ChatGPT

DeepSeek最新モデル調査レポート

エグゼクティブサマリ

2026年6月14日時点で、DeepSeekの**公式に公開済みの最新モデル系統**は **DeepSeek-V4 Preview** です。公式ニュース一覧では2026年4月24日の「DeepSeek-V4 Preview Release」が最新のモデルリリースとして掲載されており、APIドキュメントでも現行モデル名は `deepseek-v4-pro` と `deepseek-v4-flash` です。従来の `deepseek-chat` と `deepseek-reasoner` は 2026年7月24日に完全廃止予定で、現時点ではそれぞれ V4-Flash の非思考モード／思考モードに互換マッピングされています。したがって、**「V4が最新か」への答えは、公式基準では「はい。ただし、まだ Preview 表記であり、安定版確定ではない」**です。 [1]

V4系は、**V4-Pro が総1.6Tパラメータ・推論時49B活性、V4-Flash が総284B・13B活性**の MoE モデルで、いずれも**100万トークンのコンテキスト長**を公式に掲げています。技術報告では、V4系は**32T超の高品質トークン**で事前学習され、長文効率を強化する hybrid attention、mHC、Muon optimizer などを導入したと説明されています。加えて、APIはOpenAI形式およびAnthropic形式に互換で、公開weightsはMITライセンスで配布されており、**API利用とセルフホストの両方が可能**です。 [2]

性能面では、DeepSeekの自己報告では**コーディング、長文、エージェント、検索補助**にかなり強く、たとえば V4-Pro-Max は LiveCodeBench 93.5、Codeforces 3206、SWE Verified 80.6、Terminal Bench 2.0 67.9 を掲げています。一方で、**独立評価はより慎重**です。NIST/CAISI は DeepSeek V4 を「これまで評価した中国モデルで最も高性能」としつつ、**フロンティアから約8か月遅れ**と評価しました。さらに Vals の公開スコアでは、DeepSeek V4 の **LegalBench は 80.32% だが順位は 59/118** であり、**法務・知財の高精度業務で“そのまま任せる”水準とは言い切れません**。 [3]

価格面の魅力は非常に大きく、**V4-Pro は現行公開価格ベースで GPT/Claude/Gemini の旗艦級モデルより大幅に安い**とみられます。ただし、知財実務で重要なのは単価そのものよりも、**秘密情報をどこに出すか、出力に虚偽引用が混ざらないか、検索・引用・人手審査が業務設計に組み込まれているか**です。DeepSeek自身も、出力は参考情報にすぎず、医療・法務・金融等の専門判断の基礎にすべきでないと明記しています。知財業務の観点では、**最も有望なのは「リサーチ初動」「大量文書の構造化」「候補抽出」「ドラフト生成」**であり、**最終判断・法律意見・侵害認定・FTO結論の自動化**には不向きです。 [4]

結論だけ先に言えば、**DeepSeek V4 は知財実務で“導入検討に値する”が、“単独で信用してよい”モデルではない**、というのが実務的な評価です。**機密性が低い前工程**では非常に有力です。**機密性・法的影響が高い工程**では、セルフホスト化、RAG接続、出典強制、二重レビュー、モデル比較審査を前提に使うべきです。 [5]

最新モデルの位置づけと公式リリース

公式ソースを基準に整理すると、DeepSeekのモデル更新の流れは **V3 → R1 系 → V3.1 → V3.2 → V4 Preview** で、公式ニュース一覧の最新モデルリリースは **2026年4月24日公開の DeepSeek-V4 Preview** です。したがって、現時点で**V4より新しい V4.1 / V5 の公式リリースは確認できません**。ただし、名称が “Preview” のままである点は重要で、**ベンダー自身がまだ変動余地のある状態として扱っている**と読むのが妥当です。 [6]

API運用上の実務ポイントとしては、現行の公式ドキュメントが **`deepseek-v4-pro` と `deepseek-v4-flash` を最新APIモデル**として掲げる一方、旧エンドポイント名である `deepseek-chat` と `deepseek-reasoner` を

2026年7月24日で廃止予定としていることです。すでに互換目的で V4-Flash の非思考／思考モードに結び付けられているため、**新規導入やPoCでは旧名を使わず、最初から V4系の明示モデル名で組むべき**です。 [7]

ご指定のUravation記事も、2026年6月10日更新版で **DeepSeek V4 を現行の主要論点**として整理しており、日本語での実務導入論点としては参考になります。ただし、**最新モデルかどうかの判定と仕様の最終確認は、必ず公式ニュース・公式API文書を優先**すべきです。 [8]

技術仕様比較表

項目	DeepSeek-V4-Pro	DeepSeek-V4-Flash	出典・備考
公式公開日・状態	2026-04-24 / Preview	2026-04-24 / Preview	公式リリース。公式ニュース一覧でもV4が最新。 [9]
APIモデル名	`deepseek-v4-pro`	`deepseek-v4-flash`	OpenAI形式・Anthropic形式の両インターフェースで利用可能。 [10]
総パラメータ数	1.6T	284B	公式技術報告。 [11]
推論時アクティブ	49B	13B	公式技術報告。 [11]
事前学習トークン	33T	32T	公式技術報告。 [12]
学習データ概要	公開データ＋ライセンスデータ。数理・コード・Web・長文文書・論文・技術報告など	同左	DeepSeekのモデル説明では public data / licensed data を明示。V4報告では数理・コード・Web・長文・学術資料を含むと説明。 [13]
コンテキスト長	1M tokens	1M tokens	公式リリース・技術報告。 [14]
推論モード	Non-think / Think High / Think Max	Non-think / Think High / Think Max	公式モデルカード。 [15]
モダリティ	公開weights/APIの主要記述ではテキスト中心。画像入出力は未指定	同左	公開モデルは “Text Generation”。第三者評価でも text-only と整理。 [16]
代表ベンチ	MMLU-Pro 87.5、GPQA 90.1、LiveCodeBench 93.5、SWE Verified 80.6	MMLU-Pro 86.2、GPQA 88.1、LiveCodeBench 91.6、SWE Verified 79.0	DeepSeek自己報告。 [17]

項目	DeepSeek-V4-Pro	DeepSeek-V4-Flash	出典・備考
長文効率	1M文脈で、V3.2比で単一トークン推論FLOPs 27%、KV cache 10%	V4系として長文効率を強調。個別要約数値は本調査範囲では未指定	V4-Proについて公式技術報告で明示。[18]
API互換性	OpenAI SDK / Anthropic SDK互換	同左	公式API docs。[19]
ローカル実行	可能	可能	Hugging Face でweights公開、vLLM / SGLang / Docker Model Runner 利用案内あり。[20]
ライセンス	MIT	MIT	公式モデル説明・技術原理説明。商用利用可。[21]
セキュリティ・データ保持	公式サービスでは学習オプトアウト可、チャット履歴削除可、個人データは中国で処理・保管、保持期間は「必要な限り」。API専用の詳細TTLは現行docsで明示確認できず	同左	DeepSeek Privacy Policy、Terms、Context Caching。[22]

技術仕様と性能評価

V4系の技術的な焦点は、単なる“巨大化”ではなく、**長文コンテキスト効率と推論の費用対効果**にあります。公式技術報告では、V4系は hybrid attention（CSA/HCA）、mHC、Muon optimizer を導入し、**100万トークン文脈でも計算量とKV cacheを抑える**ことを狙っています。知財業務で見ると、これは**先行技術資料束、審査経過、契約束、証拠束の同時読解**にそのまま効く設計思想です。ロングコンテキストは、単に長いPDFを読めるというより、**複数文書間の整合・矛盾・差分を一度に比較しやすい**という点で価値があります。[23]

一方で、**マルチモーダル性はV4の公開仕様では強く打ち出されていません**。DeepSeekのサービス規約は一般論として「テキスト、画像、ファイル等の入力」を扱う生成AIサービスを説明していますが、V4-Pro/V4-Flash の公開モデルカードや第三者モデルディレクトリでは、**少なくとも現行の公開weights/APIモデルは text-only として整理**されています。したがって、商標ロゴ画像の類否判定や意匠図面の視覚比較のような案件で、**V4単独を“画像ネイティブ”モデルと見なすのは危険**です。そこは **Gemini系や別の視覚モデルとの組み合わせ**を前提に考えるべきです。[24]

性能の読み方で重要なのは、**DeepSeek自己報告**と**独立評価**を分けることです。DeepSeek自己報告では、V4-Pro-Max は GPT-5.4 xHigh や Gemini-3.1-Pro High と比較して、**SimpleQA-Verified、LiveCodeBench、Codeforces、Apex Shortlist** などで強い一方、**GPQA Diamond や HLE では劣後**しています。つまり、V4は「万能に最強」ではなく、**コード・長文・実行系ワークフローに比較的強く、純粋な学術推論や最高難度の広汎知識ではまだ差が残る**、というプロファイルです。知財実務に引き直すと、**調査設計、比較表、ドラフト、反復修正**には向くが、**独立した法解釈の最終判断**には向きにくい、という評価になります。[17]

検索・IR寄りの性能については、公開されているV4評価の多くが**内部評価**です。技術報告では、DeepSeekのチャットにおいて非思考モードはRAG、思考モードはagentic searchを採用し、**agentic search が RAG を 61.7% 対 18.3% で上回った**としています。また V4-Pro は V3.2 より検索Q&Aで大きく改善したとされます。ただし、これは **BEIR のような公開標準IRベンチではなく、DeepSeek自身のSearch Q&A評価**です。した

がって、先行技術調査や侵害調査にそのまま一般化するのではなく、**自社の検索基盤と権利DBを繋いだ上での再評価が必須**です。 25

独立評価では、まず Artificial Analysis が **DeepSeek V4 Pro (Max) の Intelligence Index を 52** とし、DeepSeek内では最上位、かつ出力速度も高い水準と評価しています。他方、NIST/CAISI は DeepSeek V4 を「これまでの中国モデルで最も高性能」としながらも、**フロンティア最先端より約8か月遅れ**と判断しました。さらに、Vals では **DeepSeek V4 の LegalBench が 80.32%、Vals Index が 55.62%、Latency が 1319.05s、Cost/Test が $0.83** とされており、法律系ベンチで“弱い”わけではないものの、**法務特化の観点で飛び抜けているとも言い難い**ことが分かります。 26

知財により近い公開ベンチとしては、2025年の **IPBench** が参考になります。この研究では、**DeepSeek-V3 が overall 75.8% で首位**だった一方、著者らは「現行LLMは依然としてIPタスクを信頼して扱うには不十分」と結論づけています。加えて、法的推論を評価した別研究では、**DeepSeek-R1 は中国語法務タスクで強いが、古い法知識、法的解釈の限界、事実幻覚**が主な失敗要因とされています。V4そのものの公開IPBenchスコアは本調査範囲では確認できませんでしたが、少なくとも DeepSeek系の法務・知財適性は「有望だが未成熟」という見方が安全です。 27

推論レイテンシとスループットは、**モデルそのものより、どの提供基盤で動かすか**にかなり左右されます。独立ベンチによると、V4-Pro(Max) の**公式DeepSeek API**では、**TTFT 1.85秒、time to first answer token 128.46秒、出力速度 34.6 t/s** という結果が出ています。一方で Fireworks など他の提供基盤では time to first answer token が大きく短縮されています。つまり、**高思考モードの V4-Pro は“賢いが待たせる”モデル**であり、知財部門のバッチ型レビューには向く一方、クライアント対話の即応用途にはそのままでは使いづらいことがあります。 28

評判・料金・ライセンス

第三者レビューとユーザー評判を見ると、DeepSeekは**「価格破壊力」と「性能の伸び」を評価される一方、企業導入ではガバナンス懸念が大きい**、という二面性があります。Google Play 上の公式アプリは **4.1星、約30万件レビュー、5000万超ダウンロード**で、ユーザーレビューには「無料でかなり使える」「説明が丁寧」といった好意的評価も見られます。一方で、企業向け論評では、**プライバシー保証、コンプライアンス、監査性の不足が導入障壁**として繰り返し指摘されています。 29

企業導入事例については、**V4固有の一次事例はまだ少ない**です。本調査で公表確認できたのは、主に**DeepSeekファミリー全体**としての導入例で、Reuters は **Tiger Brokers が DeepSeek モデルをチャットボットに統合し、20社超の中国証券・資産運用会社がDeepSeek系を活用**していると報じています。ただし、これは主としてR1以降の DeepSeek採用の広がりを示すものであり、**V4-Pro/V4-Flash 単体の成熟したエンタープライズ事例集が揃っている段階ではありません**。この点は導入判断で重要です。 30

料金・ライセンス比較表

製品	入力価格	キャッシュ入力	出力価格	ライセンス / 配備	データ利用・保持の要点	出典
DeepSeek V4-Pro	**$0.435 / 1M**	**$0.0036 / 1M**	**$0.87 / 1M**	MIT / API + 公開weights + セルフホスト可	個人データの学習オプトアウト可。チャット履歴削除可。個人データは中国で処理・保管。保持期間は「必要な限り」。API専用の詳細TTLは現行docsで明示確認できず。	31
DeepSeek V4-Flash	**$0.14 / 1M**	未確認	**$0.28 / 1M**	MIT / API + 公開weights + セルフホスト可	上記DeepSeek共通方針。Flashのキャッシュ単価は本調査で明示確認できず。	32
OpenAI GPT-5.5	**$5.00 / 1M**	**$0.50 / 1M**	**$30.00 / 1M**	Proprietary / API（ホスト型）	API送信データはデフォルトで学習不使用。乱用監視ログは通常最長30日。承認制でZero Data Retention等あり。	33

製品	入力価格	キャッシュ入力	出力価格	ライセンス / 配備	データ利用・保持の要点	出典
Anthropic Claude Opus 4.8	**$5.00 / 1M**	**$0.50 / 1M**（cache hit）	**$25.00 / 1M**	Proprietary / API（直販・Bedrock等）	モデル学習には明示許可が必要。標準では会話内容はデフォルト保持なし。ZDRとHIPAA-readyあり。	34
Google Gemini 3.1 Pro Preview	**$2.00 / 1M**（<=200k）	**$0.20 / 1M**（<=200k）	**$12.00 / 1M**（<=200k）	Proprietary / API（AI Studio / Cloud）	Paid Servicesではプロンプト・応答は製品改善に不使用。ZDRあり。ただしGrounding with Google Search / Maps では30日保存が発生。	35

上表だけを見ると、**DeepSeek V4-Pro の価格競争力はかなり強い**です。V4-Pro は GPT-5.5 の 1/10 以下、Claude Opus 4.8 の 1/5〜1/30 レベルの単価帯で、しかも公開weightsがあるため、**“SaaSで試す→セルフホストへ移行” のオプション価値**まで持っています。ただし、ここで見落としがちなのが**出力トークン量**です。Artificial Analysis は V4-Pro(Max) について、Intelligence Index 評価で **190M output tokens** を生成し、同規模オープンモデルよりかなり冗長だと指摘しています。つまり、**単価が安くても、長考・長文出力を許す運用では請求総額が膨らみやすい**ため、知財業務では「要約長制御」「理由の深さの段階制」「引用必須時のみ Max」などのポリシーが必要です。 36

知財業務での実務ユースケース

知財業務における DeepSeek V4 の使いどころは、**法的結論を自動化することではなく、文書・証拠・検索候補・比較観点を高速に整えること**です。V4系は、公式技術報告ベースでは**検索強化、agentic search、ホワイトカラー文書生成、13業界をまたぐ高度タスク**を重視しており、その中には law も含まれています。ただし、公開の LegalBench・IPBench・法的推論研究は、**法務ドメインでまだ無視できない幻覚・誤引用・法解釈限界がある**ことを示しています。したがって、以下のユースケースはすべて、**権利DB・判例DB・社内DMS・契約台帳への接続**と、**人間の最終承認**を前提に設計するのが実務的です。 37

知財ユースケース一覧表

分野	ユースケース	推奨ワークフロー	期待効果	精度要件・検証方法	主なリスク
特許	発明ヒアリング整理	Flashで面談メモ要約 → Proで発明要素・課題・効果・代替案を構造化	受付品質の平準化	必須項目欠落率を人手点検、抜け漏れ率を月次監査	事実補完の幻覚、秘匿事項漏えい
特許	クレーム用語正規化	仕様書・発明メモから用語表作成 → 用語揺れを一覧化	用語統一、明細書品質向上	用語集一致率 95% 以上を目標	誤同義語化、意味の狭窄
特許	CPC/IPC候補提示	要約から分類候補を複数提示 → 審査官・担当者が選択	初動短縮	既知案件でTop3再現率を測定	誤分類による検索漏れ
特許	先行技術検索式拡張	発明要素→同義語・上位概念・用途語を生成 → DB検索に投入	検索網羅性向上	既存ゴールドセットで Recall@k を確認	不要語混入、ノイズ増大
特許	先行文献スクリーニング	Top N文献を要約して請求項との差分一覧を作成	読み込み負荷削減	人手ラベルで relevant / irrelevant 一致率検証	関連文献の見落とし
特許	新規性ホットスポット抽出	独立請求項の各構成要件ごとに先行例をマッピング	差別化論点の早期把握	構成要件単位で人手照合	要件対応の捏造
特許	クレームチャート下書き	文献引用箇所付きで claim chart 草案生成	起案時間短縮	**引用URL/段落/図番号の実在率100%** を必須	架空引用、誤マッピング
特許	OA論点抽出	拒絶理由通知書から引用文献・拒絶理由・応答案構成を抽出	応答案の立ち上がり高速化	過去案件で issue extraction F1 を測定	理由誤読、法的結論の先走り
特許	OA応答骨子案	OA論点と明細書から、反論・補正・代替案の骨子を生成	思考のたたき台作成	弁理士レビュー採用率を追跡	危険な補正案、禁反言リスク見落とし
特許	ファミリー・法的状態整理	各国出願・審査・登録・失効情報を時系列表に整形	多国管理の見通し改善	Docket台帳との一致確認	ステータス誤同期
特許	IDS候補重複排除	文献候補を family / pub / assignee 単位で dedup	重複工数削減	人手レビューで過剰排除率を計測	重要異本の誤排除

分野	ユースケース	推奨ワークフロー	期待効果	精度要件・検証方法	主なリスク
特許	FTO製品仕様抽出	製品仕様書・マニュアルから比較に必要な技術要素を抽出	FTO前処理高速化	抽出項目の再現率検証	非公開仕様のクラウド流出
特許	規格×特許対応表	標準文書と請求項の候補対応箇所を一覧化	SEP分析の初動短縮	引用箇所の実在率・再現率を監査	過剰マッチング
特許	出願書類整合性レビュー	明細書・クレーム・要約・図面説明の矛盾点を抽出	品質検査自動化	矛盾検出の再現率を案件別に記録	擬陽性過多
商標	クリアランス検索語生成	商標候補から称呼・観念・表記揺れ・翻字を生成	検索網羅性向上	既知NG案件で再現率確認	重要類似候補の漏れ
商標	類似群コード候補整理	指定商品役務から候補区分・類似群を下書き	分類初動短縮	過去出願との一致率確認	区分誤り
商標	指定商品役務文案作成	事業内容から審査実務に沿った文案草案作成	文案作成効率化	審査補正率を継続測定	過度に広い/狭い指定
商標	称呼・外観・観念の三面比較	候補商標群の比較表を作成	判断材料の可視化	人手レビュー前提、結論自動化禁止	類否結論の飛躍
商標	ウォッチ通知のトリアージ	監視結果を重要度別に並べ、対応候補を提案	通知処理の優先付け	拒否/保留/要確認の精度を測定	重要案件の取りこぼし
商標	異議・無効骨子案	登録商標と先行権利を対比し論点草案を作成	起案スピード向上	人手の全面レビュー必須	誤法域・誤条文
商標	使用証拠チェックリスト	業務実態から必要証拠一覧を生成	証拠収集漏れ防止	必須証拠の欠落率監査	国別証拠要件誤り
商標	海外出願訳語統制	JP/EN/CN の商品役務表現を用語集準拠で整形	多言語整合性向上	用語集一致率測定	誤訳で権利範囲変質
著作権	侵害主張比較表	原著作物と対象物の構成要素比較表を作成	争点の見える化	根拠資料リンク必須、結論は人手	類似性の誤評価
著作権	ソースコード由来調査補助	OSS断片・コメント・ヘッダから候補ライセンスを抽出	調査の入口形成	SPDX/既知ライセンスで正答率確認	誤ライセンス判定
著作権	OSSライセンス条項抽出	依存関係一覧→主要義務・表示要件・copyleft整理	契約・配布準備効率化	重要義務の見落とし率を点検	義務漏れ

分野	ユースケース	推奨ワークフロー	期待効果	精度要件・検証方法	主なリスク
著作権	学習データ/権利処理質問票作成	生成AI案件向けにデータ出所・許諾・除外手続を質問票化	初回DD効率化	質問票網羅率を法務確認	重要論点の未聴取
著作権	出版・Webコンテンツ権利表	著作者、利用許諾、二次利用可否を一覧化	権利台帳整備	既存契約台帳との一致確認	古い契約引用
著作権	削除要請/反論骨子	プラットフォーム通知文、DMCA系文面の草案補助	初動短縮	事実誤認率をレビュー	断定表現による紛争悪化
契約	AI/知財条項レビュー	ProでIP帰属・利用範囲・再学習・監査条項を抽出	論点見落とし減少	Issue spotting recall を過去案件で測定	条文誤読
契約	NDAの秘密情報条項点検	例外、残存義務、返還・削除条項を抽出	レビューの標準化	条項抽出F1 / 人手再確認	秘密情報定義の誤解
契約	職務発明・発明譲渡監査	雇用契約・就業規則・覚書から権利帰属差異抽出	ルール不整合の発見	抽出結果を労務・知財で二重確認	規程改訂漏れ
契約	共同研究契約の境界論点整理	バックグラウンドIP、改良発明、公表、実施権を比較	交渉論点の可視化	重要論点漏れ率を点検	結論の単純化
契約	SaaS/DPAのデータ越境確認	保管地、下請先、監査権、削除対応を抽出	ベンダー審査迅速化	標準チェックリストと突合	規制要件の取り違え
契約	補償・保証条項レビュー	IP indemnity、免責上限、除外条項を要約比較	リスク把握迅速化	ハイリスク条項の抽出率確認	単語一致だけで見落とす
横断	証拠年表・ファクトマトリクス	メール、議事録、版管理履歴を時系列整理	紛争準備の初速向上	証拠番号リンク必須	証拠誤接続
横断	多言語調査レポート草案	JP/EN/CN資料を統合し出典付き要旨作成	国際案件の情報統合	用語集・引用整合テスト	誤訳、出典捏造
横断	案件メールの振分け	OA、ウォッチ、契約依頼、侵害相談を分類	受付効率化	分類精度 / 誤振分率	重要案件の誤分類
横断	レビューコメント要約	共同レビューのコメント束を論点別に整理	合意形成の高速化	人手差分比較	コメントの意味落ち

これらのユースケース設計の根拠は、V4系の**長文処理、検索・agentic search、文書生成、ホワイトカラー・コード能力**にあります。ただし、公開の LegalBench / IPBench / 法的推論研究、および DeepSeek 自身の免責表示が示すとおり、**“結論”ではなく“候補・整理・比較・下書き”に留めること**が大前提です。 38

法的・倫理的リスク

最も重いリスクは、**機密情報・営業秘密・依頼者特権情報の取り扱い**です。DeepSeek の Privacy Policy では、個人データは**中国で直接収集・処理・保存**されうると明記され、保持期間はサービス提供や法的義務等に必要な限りとされています。学習オプトアウトやチャット履歴削除は可能ですが、**知財実務で重要なのは“削除できること”より“最初から外に出さないこと”**です。少なくとも、特許出願前の発明内容、侵害判断メモ、和解方針、ライセンス交渉ドラフトを**公式Web/Appにそのまま投入するのは避けるべき**です。 39

品質面では、DeepSeek自身が**幻覚の可能性を否定しておらず、法務・医療・金融等で専門助言とみなすべきでない**と明示しています。さらに、法的推論研究では、DeepSeek系に**古い法知識、法解釈の弱さ、事実幻覚**があることが示されています。このため、**実在しない判例、存在しない文献、条文の取り違え**は、知財業務で実害に直結します。特に**先行技術調査・商標調査・侵害比較**では、**“引用元の存在確認が100%取れない出力は利用禁止”**という運用ルールが必要です。 40

知財権との関係では、DeepSeek の Terms は、ユーザーが入力した Inputs の権利を保持し、**Outputs の権利を“ある場合に限り”ユーザーへ譲渡する**としています。他方で、**出力はユニークではない可能性**があり、他ユーザーに類似出力が生成されうることも明記されています。つまり、生成ドラフトをそのまま「自社独自の表現」と扱うのは危険です。また、Terms はユーザーが Outputs を**他モデルの学習や蒸留にも用いうる**としていますが、これは逆にいえば、**出力が第三者権利を侵害しない保証をベンダーが与えていない**ことも意味します。知財文書では、**類似度チェック、出典確認、権利帰属確認**が不可欠です。 41

地政学・規制リスクも無視できません。DeepSeek の Terms は準拠法を**中国本土法**、裁判管轄を**杭州深度求索の所在地法院**としています。加えて、Reuters や AP は、個人情報保護や政府アクセス懸念を理由に、**韓国当局の問題提起やチェコ政府の禁止措置**を報じています。OpenAI・Anthropic 側からは DeepSeek への蒸留・不正アクセスを巡る疑義も公に提起されており、これはまだ**係争や行政判断で確定した事実ではない**ものの、**取締役会・情報セキュリティ委員会での説明責任**を要するリスクです。 42

最後に、**Preview版であること自体が運用リスク**です。モデル挙動や価格が変わる可能性があり、旧API名の廃止予定も既に告知されています。知財実務で使うなら、**モデル名の固定、評価セットによる回帰試験、プロンプトバージョン管理、エスカレーション先の明確化**が必須です。 7

導入ガイドと結論

知財部門で DeepSeek V4 を試すなら、PoC は**“一般チャットがどれだけ滑らかか”ではなく、“既知案件でどれだけ事故なく再現できるか”**で設計すべきです。NIST/CAISI が示すように、ベンダー自己評価と独立評価には差があり、公開 LegalBench でも V4 は中位帯です。したがって、PoC の単位は「会話満足度」ではなく、**案件タイプ別の再現性と監査可能性**に置くべきです。 43

実務的なPoC設計は、まず **三つの業務束** から始めるのが合理的です。第一に、**検索・抽出系**（先行技術候補抽出、商標候補整理、契約条項抽出）。第二に、**比較・整理系**（claim chart 下書き、類否比較表、証拠年表）。第三に、**ドラフト系**（OA応答骨子、契約レビューコメント、質問票）。各束で**ゴールドセットを50〜100件程度**作り、既知の正解または少なくとも「採否判断済みの過去成果物」を持つデータだけで評価します。ゴールドセットは**公開資料版**と**秘密情報含有版**に分け、後者はセルフホスト環境でのみ評価するのが安全です。 44

評価指標は、知財実務では最低でも次の四群が必要です。**正確性**としては citation precision、条項抽出F1、検索 Recall@k。**業務適合性**としては reviewer acceptance rate と issue spotting recall。**安全性**としては unsupported assertion rate、機微情報漏えい件数、誤引用率。**運用性**としては 1件あたり処理時間、トークンコスト、再実行率です。特に、**侵害調査・FTO・法的結論**に近づくほど、目標は「高精度」ではなく**“誤っ**

た断定をしないこと”に置くべきです。つまり、**自動完了率**より**保守的な要確認率**を重視する評価が適しています。[45]

データ準備では、モデルに直接“考えさせる”前に、**権威ソースを先に整備**するのが成功条件です。具体的には、特許なら J-PlatPat / Espacenet / USPTO / WIPO、商標なら JPO / USPTO / EUIPO / TMview、著作権・OSSなら契約台帳・レジストリ・リポジトリ台帳、契約なら DMS と clause library を用意します。DeepSeek V4 は**検索と文書処理に強い**ので、最も効果が出るのは「DBアクセス→候補束→モデル整理」という順序です。**モデル単独検索に期待しすぎる構成**は避けてください。[46]

運用体制としては、**モデルルーティング**が有効です。私なら、**Flash** を「大量文書の抽出・分類・候補生成」、**Pro High** を「比較分析・起案」、**Pro Max** を「高難度の論点整理や第二案作成」に限定します。そして、**最終レビュー前に別系統モデルでクロスチェック**する運用を推奨します。理由は、DeepSeek V4 は価格優位が大きい一方、独立評価では最前線の閉鎖モデルにまだ差があり、また法務ベンチでは上位独占ではないからです。**DeepSeek を“主査”ではなく“副査・下書き担当”として置く**のが、現時点では最も失敗しにくい構えです。[47]

競合比較表

製品	デプロイ / ライセンス	モダリティ	検索・ツール	強み	弱み・留意点	知財業務での推奨ポジション	出典
DeepSeek V4-Pro	API + 公開 weights / MIT	テキスト中心	DeepSeek ChatではRAG と agentic search。API はOpenAI/ Anthropic互換。	1M文脈、低価格、セルフホスト可能、コード・長文・比較表に強い	Preview、text-only寄り、SaaSは中国保管・準拠法、独立評価では最前線に未達	**低〜中リスクの下書き・整理・検索補助の主力**	[48]
OpenAI GPT-5.5	API / Proprietary	テキスト+画像入力配慮 + 多数のホストツール	Web search、file search、computer use、code interpreter 等	ツール群が豊富。高難度の調査・反復ワークフロー設計に強い	高価格、セルフホスト不可	**最終クロスチェック、複雑調査、エージェント統合作業**	[49]

製品	デプロイ / ライセンス	モダリティ	検索・ツール	強み	弱み・留意点	知財業務での推奨ポジション	出典
Anthropic Claude Opus 4.8	API / Proprietary	長文テキスト中心	Web search、web fetch、Managed Agents。ZDR/HIPAA対応あり	1M文脈、長文読解と監査性が強い。エンタープライズ運用の安心感が高い	高価格、セルフホスト不可	**高信頼の契約レビュー・証拠束読解・監査重視の用途**	[50]
Google Gemini 3.1 Pro Preview	API / Proprietary	マルチモーダル	Grounding with Google Search / Maps、カスタムツール	マルチモーダル、検索グラウンディング、Paidで学習不使用	Preview、Grounding利用時は30日保存、価格はDeepSeekより高い	**商標画像・Web証拠・多媒体証拠束の統合評価**	[51]

コスト見積の目安

DeepSeek V4-Pro の月額概算は、`0.435 × 入力MTok + 0.87 × 出力MTok + 0.0036 × キャッシュMTok` で近似できます。V4-Flash は公開二次情報ベースで `0.14 × 入力MTok + 0.28 × 出力MTok` を最低線の目安にできますが、**キャッシュ単価は本調査で明示確認できていません**。[52]

シナリオ	月間入力	月間出力	月間キャッシュ	V4-Pro 概算	V4-Flash 概算
小規模PoC	20 MTok	5 MTok	100 MTok	**約 $13.41**	**約 $4.20** 以上
部門運用	200 MTok	50 MTok	1,000 MTok	**約 $134.10**	**約 $42.00** 以上
全社バッチ処理	2,000 MTok	500 MTok	10,000 MTok	**約 $1,341.00**	**約 $420.00** 以上

同じ「部門運用」ボリュームを OpenAI GPT-5.4 の標準単価で計算すると、**入力 $500 + 出力 $750 + キャッシュ $250 = 約 $1,500** となり、V4-Pro の約 $134 に対してかなり大きい差が出ます。したがって、**大量文書を読む知財バックオフィス用途では DeepSeek の経済合理性は非常に高い**と言えます。ただし、その差額を**二重レビュー、引用検証、セルフホスト基盤**に再投資して初めて、実務的な優位になります。[53]

最終的な推奨は明確です。**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、“高コスパの実務補助モデル”として導入するのが最適**です。具体的には、**調査初動、論点整理、比較表、ドラフト、台帳化、翻訳整形**では積極活用してよい。一方で、**FTO結論、侵害認定、最終意見書、クライアント提出版の断定表現**は、人間レビューに加えて別系統モデルまたは権威ソース照合を通すべきです。**機密性が高い**

案件ではセルフホスト前提、SaaS利用時は**中国保管・準拠法・保持方針を許容できる案件に限定**、これが2026年6月時点の最も実務的な結論です。 54

navlist関連ニュースと政策動向
turn12news44,turn9news29,turn22news38,turn21news37,turn22news39

---

1 6 https://api-docs.deepseek.com/news/news251201
https://api-docs.deepseek.com/news/news251201

2 11 12 18 23 25 37 38 46 https://huggingface.co/deepseek-ai/DeepSeek-V4-Pro/resolve/main/DeepSeek_V4.pdf?download=true
https://huggingface.co/deepseek-ai/DeepSeek-V4-Pro/resolve/main/DeepSeek_V4.pdf?download=true

3 15 16 17 https://huggingface.co/deepseek-ai/DeepSeek-V4-Pro
https://huggingface.co/deepseek-ai/DeepSeek-V4-Pro

4 5 13 21 40 https://cdn.deepseek.com/policies/en-US/model-algorithm-disclosure.html
https://cdn.deepseek.com/policies/en-US/model-algorithm-disclosure.html

7 10 19 48 https://api-docs.deepseek.com/
https://api-docs.deepseek.com/

8 32 https://uravation.com/media/deepseek-v4-preview-complete-guide-2026/
https://uravation.com/media/deepseek-v4-preview-complete-guide-2026/

9 14 https://api-docs.deepseek.com/news/news260424
https://api-docs.deepseek.com/news/news260424

20 https://huggingface.co/deepseek-ai/DeepSeek-V3.2
https://huggingface.co/deepseek-ai/DeepSeek-V3.2

22 39 44 54 https://cdn.deepseek.com/policies/en-US/deepseek-privacy-policy.html
https://cdn.deepseek.com/policies/en-US/deepseek-privacy-policy.html

24 41 42 https://cdn.deepseek.com/policies/en-US/deepseek-terms-of-use.html
https://cdn.deepseek.com/policies/en-US/deepseek-terms-of-use.html

26 31 36 52 https://artificialanalysis.ai/models/deepseek-v4-pro
https://artificialanalysis.ai/models/deepseek-v4-pro

27 https://arxiv.org/html/2504.15524v2
https://arxiv.org/html/2504.15524v2

28 https://deepinfra.com/blog/deepseek-v4-pro-max-api-benchmarks-latency-throughput-cost
https://deepinfra.com/blog/deepseek-v4-pro-max-api-benchmarks-latency-throughput-cost

29 https://play.google.com/store/apps/details?hl=en&id=com.deepseek.chat
https://play.google.com/store/apps/details?hl=en&id=com.deepseek.chat

30 https://www.reuters.com/technology/artificial-intelligence/tiger-brokers-adopts-deepseek-model-chinese-brokerages-funds-rush-embrace-ai-2025-02-18/
https://www.reuters.com/technology/artificial-intelligence/tiger-brokers-adopts-deepseek-model-chinese-brokerages-funds-rush-embrace-ai-2025-02-18/

33 49 53 https://openai.com/api/pricing/
https://openai.com/api/pricing/

34 50 https://docs.anthropic.com/en/docs/about-claude/models
https://docs.anthropic.com/en/docs/about-claude/models

35 51 https://ai.google.dev/gemini-api/docs/pricing
https://ai.google.dev/gemini-api/docs/pricing

43 47 https://www.nist.gov/news-events/news/2026/05/caisi-evaluation-deepseek-v4-pro
https://www.nist.gov/news-events/news/2026/05/caisi-evaluation-deepseek-v4-pro

45 https://arxiv.org/html/2503.16040v1
https://arxiv.org/html/2503.16040v1