ChatGPT

Google「Deep Review」の知財業務活用に関する調査・導入分析

エグゼクティブサマリ

Google Researchは2026年10月14日、長文の技術文書を読み、論理的矛盾、数学的・技術的な誤り、根拠不足、既存研究との不整合などを指摘するエージェント型AI「Deep Review」を一般公開した。CNET Japanも2026年10月16日付で「論文の矛盾を見つけるAI『Deep Review』、Googleが公開　週2回まで無料」と報じている。CNET記事のYahoo!ニュース版は本調査環境から本文を直接取得できなかったため、同一CNET Japan記事の配信版と、Google Researchの一次発表を突合して確認した。 [1]

知財業務との相性は高いが、「先行技術検索AI」ではなく「技術・論理レビューAI」と位置付けるべきである。 Deep Reviewは文書を論理単位に分割し、難易度に応じて計算量を配分し、複数の推論エージェントが並列検証した後、全体を再統合して外部情報によるgroundingまで行う。この構造は、特許明細書の「請求項―実施形態―実施例」の整合確認、拒絶理由通知に対する反論の弱点検出、無効資料・クレームチャートのストレステスト、技術DDやライセンス資料の根拠確認などに適する。Google自身も法律・金融文書について、条項・表間の整合性、曖昧な用語、下流への影響、根拠確認を用途として明示している。 [2]

一方、**Deep Review単独で先行技術調査、特許性判断、FTO、無効判断を完結させるべきではない。** Googleが公開している精度指標は特許案件ではなく、数学・コンピュータ科学論文の「式／証明エラー」に絞ったSPOTベンチマーク26論文・29エラーに対する89.7%であり、しかも原SPOTとは採点方法が異なるため数値を直接比較できない。元のSPOTでは最先端LLMでも83論文・91エラー全体で最大21.1% recall、6.1% precisionに留まっており、「高性能なLLMでも検証タスクは難しい」という前提は維持すべきである。 [3]

導入上の最大の論点は機密性である。 Deep ReviewのWeb UIについてGoogleは、アップロード文書をモデルの学習・微調整・改善には使用しない一方、UI表示のため**1年間保存する**と明示している。これは学習利用リスクを大きく下げるが、未公開発明、営業秘密、クライアント秘密を第三者クラウドへ預ける問題とは別である。日本弁理士会は、弁理士の守秘義務、NDA、新規性、個人情報の観点から外部生成AIへの秘密情報入力に慎重な判断を求め、特にクライアント秘密を入力する場合には「学習の有無によらずクライアントに合意を取るべき」としている。 [4]

さらに、**Deep Review UIとGemini API経由の利用条件を混同してはならない。** 一般のGemini API規約では、無料API枠などの「Unpaid Services」は入力・出力をGoogle製品改善に利用し得て、人間のレビュー対象にもなり得るため、Google自身が秘密・機密・個人情報を入力しないよう求めている。一方、課金済みCloud Projectを介した「Paid Services」では、プロンプト／レスポンスを製品改善に使わないと規定されている。ただし、**Deep Review API固有の料金、エンドポイント、SLA、保持期間、データ所在地、Deep Review UIの特別ポリシーがAPIにもそのまま適用されるかは、2026年10月16日時点の公式発表では未指定**である。 [5]

したがって、本調査の推奨方針は次の通りである。

判断	推奨
導入可否	**PoC導入を推奨**

判断	推奨
最初の用途	公開済み特許・過去案件を用いた「第二査読者」
最も有望	出願書類QA、OA対応の反論チェック、無効・訴訟準備のストレステスト
先行技術調査	検索そのものではなく、検索結果・検索戦略・クレームチャートの検証に使用
未公開発明	Web無料版への直接投入は原則避ける。契約・クライアント同意・APIデータ条件確認後に拡大
自動判断	不可。特許性・侵害・無効等の最終判断は弁理士・弁護士・技術専門家
技術構成	短期はSaaS、実務化は**特許DB＋Deep Review API＋人間レビューのハイブリッド**
精度評価	Google論文の89.7%を自社精度とみなさず、30〜50件以上の過去案件で独自ベンチマーク
ROI	標準シナリオでは月60レビュー、0.75時間/件削減で月36万円の工数価値。仮定API/運用費を引くと月14万円、初期120万円なら約8.6か月で回収という試算
最重要KPI	validated precision、重大指摘recall、レビュー時間、誤検出確認時間、見逃し率、機密事故ゼロ

調査項目と結論の一覧

調査項目	本調査の結論
機能・制限	PDF／プレーンテキスト対応、複数ファイルは今後。週2レビュー無料。APIあり。詳細API仕様・専用料金は未指定
技術的背景	分割→動的計算配分→並列深掘り→全体統合・groundingという長文向けmulti-agent verification
既存研究との比較	従来NLIより長文・多段推論型。単一LLMやpass@kの「見逃し／誤検出増加」を抑える設計
先行技術調査	補助用途として中〜高。検索DBの代替にはならない
出願書類作成	高適合
審査対応	高適合。ただし現状のsingle-file制約への対応が必要
無効審判・訴訟	高価値だが、機密・証拠性・privilege/守秘のリスクも最大
技術評価・ライセンス	高適合。技術根拠・前提条件のストレステストに有効
期待効果	工数削減と第二査読による品質向上。ただし知財固有の実証値は現時点で存在しない
法的留意点	守秘、新規性、個人情報、クライアント同意、AI誤検出、証拠化が主要論点
実装	UI→APIハイブリッドが現実的。Deep Reviewそのもののオンプレ版は未指定
ROI	API単価未指定のため、複数仮定による感度分析が必要

Deep Reviewの機能・技術的背景と制限

何が公開されたのか

Deep Reviewは、Google Research、Google DeepMind、Google Cloudが共同で開発した文書検証システムで、研究論文向けの実験システム「Paper Assistant Tool（PAT）」を一般化したものである。STOC 2026、ICML 2026、NeurIPS 2026、ICLR 2027向けプログラムでは、オプトインした**16,000件超の原稿**に事前レビューを提供した後、2026年10月14日に一般公開された。現在のDeep Reviewは従来PATより新しい「Gemini 4 Argon」をサポートしており、GoogleはPAT論文で報告した世代より能力が向上したとしているが、現行版固有の特許・法律分野ベンチマークはまだ公開していない。 [6]

CNET Japanは一般向けに「論文や技術資料の論理的な矛盾や根拠の不足を点検するAI」と紹介しており、Google一次情報ではさらに、論理整合性、数学的主張、改善提案、事実のgrounding、既存研究との比較までを対象としている。 [7]

公開仕様

項目	2026年10月16日時点の確認結果	知財実務への意味
入力	PDF、プレーンテキスト	明細書、意見書、技術報告書には適合
複数ファイル	**未対応、coming soon**	OA＋引用文献＋明細書の横断レビューには現状工夫が必要
出力	詳細レビュー／技術検証レポート	論点抽出・指摘リストとして利用
論理検証	対応	クレームと実施形態、反論内の矛盾検出に有望
数学・数値検証	対応	実施例、数式、閾値、表との整合確認に有効
事実grounding	対応	引用や外部主張の再確認に有用。ただし特許DBの完全性は保証されない
既存研究比較	対応	技術背景のレビューには利用可能
無料枠	**1ユーザーにつき週2文書レビュー**	公開・匿名化文書でのPoCには十分
高負荷利用	Gemini API key接続、Deep Review API	業務統合可能
Deep Review APIのendpoint/schema	**未指定**	実装前に公式仕様確認が必須
Deep Review固有API料金	**未指定**	ROIは仮定レンジで評価すべき
API SLA	**未指定**	本番ワークフローの期限保証に注意
UIアップロードの学習利用	使用しない	学習漏洩リスクは低減
UI文書保存	**1年間**	未公開案件では重要なリスク要因

項目	2026年10月16日時点の確認結果	知財実務への意味
オンプレ版	**未指定／公表なし**	Deep Reviewそのものを完全閉域化する前提は置けない

以上はGoogleの2026年10月14日発表に基づく。 [8]

「週2回」はレビュー回数であり、組織向けに人数分をプールできるとは公式発表されていない。例えば正規ユーザー10名なら数学的には月平均約86.7レビュー相当（10名×2回×52週÷12か月）になるが、これは**組織向けライセンス枠を意味するものではない**。また、無料だから未公開案件を投入してよいという意味でもない。

矛盾検出は、単純な「文章AとBのcontradiction分類」ではない

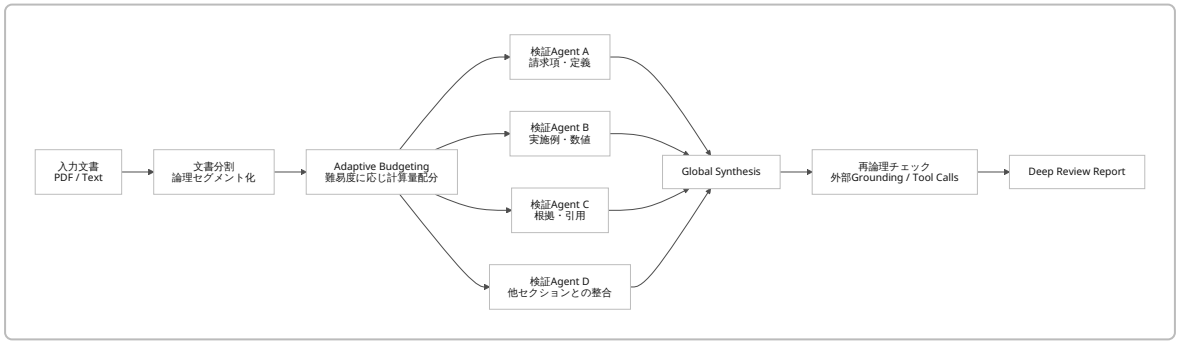
自然言語処理における矛盾検出は古くからNatural Language Inference（NLI）として研究されてきた。de Marneffe、Rafferty、Manningらは2008年に、否定、反意語、数値差、不整合する事実などを手掛かりにテキスト間の矛盾を識別する研究を発表した。MultiNLIはその後、premiseとhypothesisの関係をentailment／neutral／contradictionに分類する大規模ベンチマークを提供した。 [9]

しかし、特許明細書のような100ページ近い文書について、

> 「請求項17の限定が段落[0048]の定義と整合し、その定義が実施例3の数値と整合し、さらにOA反論中の解釈とも矛盾しないか」

を確認する仕事は、2文間のNLIとは本質的に異なる。

GoogleのDeep Reviewは次の四段階を採る。 [10]



第一に、文書を論理的セグメントへ分割する。セグメントは重複してよく、入力全体をカバーする。第二に、情報密度や複雑さによって思考計算量を動的に変える。第三に、各専門エージェントが担当部分を深く確認する一方、**全文をcontextとして保持**する。第四に、synthesis agentが指摘を統合・重複排除し、必要なら論理検証を再実行し、検索等による外部groundingを行う。 [11]

この構成は、単一プロンプトの長文レビューでは「重要な節に十分なreasoning tokenを割けない」問題、単純に独立試行を大量に行うpass@kでは「recallが上がってもfalse positiveや重複指摘が増える」問題を緩和することを狙っている。PAT論文は、セグメント別の深い推論、動的budgeting、最後のgrounding／deduplicationをその対策としている。 [12]

既存方式との位置付け

方式	単位	強み	弱み	Deep Reviewとの差
古典的矛盾検出/NLI	文・文ペア	高速、評価しやすい	長距離依存、数式、外部事実が弱い	Deep Reviewは全文構造・多段推論
単一LLMレビュー	文書全体	実装が容易	長文で注意・思考budgetが分散	セグメント化＋分業
Pass@k	同じ問いを複数試行	見逃し率低減	誤検出・重複も増加	最後にsynthesis・grounding
ACL 2025の別系統「DeepReview」	学術査読	文献検索、構造化評価、evidence-based argument	Google製品とは別プロジェクト	名称は同じだが別技術
Google Deep Review/PAT	長文	適応的計算、並列検証、全体統合、grounding	LLM誤推論、PDF解析、コスト	今回の対象

注意すべきは、2025年ACLにZhuらが発表した**同名の「DeepReview: Improving LLM-based Paper Review with Human-like Deep Thinking Process」**が存在することである。こちらはDeepReview-13Kデータセットと DeepReviewer-14Bを使い、structured analysis、literature retrieval、evidence-based argumentationを行う研究であり、**Googleの2026年Deep Reviewとは別物**である。技術思想には共通点があるものの、精度数値やライセンスを混在させてはならない。 13

精度をどう読むべきか

Google PATが使用したSPOTサブセットは、数学・CS分野の「equation/proof」エラーを含む26論文、29個の既知エラーである。結果は以下だった。 12

方法	検出率
original SPOT SOTA	21.1%
Gemini 3.1 Pro単発	55.2%
PAT + Gemini 3.1 Pro	**89.7%**

単発モデルからPATへの改善は**34.5 percentage points**である。

ただしGoogle自身が、「PAT側は論理的に同等の指摘も正解と判断する別graderを使っているため、original SPOTの21.1%とは直接比較できない」と明記している。また29エラーという小規模で、特許文書でもない。したがって、**「Deep Reviewの特許誤り検出精度89.7%」と表現するのは不適切**である。 12

現場パイロットでは、STOCの回答者124名、ICMLの733名のうち「再度使いたい」が97%／92.1%、「明瞭性・可読性が改善」が85.1%／87.0%、「very/mostly helpful」が92.7%／90.7%だった。一方、「feedback

がmostly/all grounded」と回答した比率は55.8%／64.8%であり、**人間による検証を省ける水準とは読めない**。[14]

実際、PATの既知の失敗例にはPDF parsing、古い知識・日付hallucination、そして「正しい証明・議論を誤って間違いだと断定する」reasoning failureがある。最後のタイプはLLMに固有の課題としてGoogle自身が明示している。[15]

知財業務プロセス別の適用可能性

Deep Reviewの価値は、「AIに特許判断をさせる」ことではなく、**通常は専門家一人が一度しか読まない文書に、独立した第二の技術査読レイヤーを追加すること**にある。

適用可能性マトリクス

業務	適合度	Deep Reviewが得意な役割	任せてはいけない役割
先行技術調査	★★★☆☆	発明説明の矛盾、検索観点漏れ、検索結果との技術比較、クレームチャートQA	全世界先行技術の網羅検索、日付・legal status保証
出願書類作成	★★★★★	claim/spec/example整合、定義、数値、論理飛躍、記載漏れ候補	新規性・進歩性・サポート要件等の最終法律判断
審査対応	★★★★★	OA反論のstress test、補正と原記載の対応チェック	審査官の法的評価を確定的に覆す判断
無効審判・訴訟準備	★★★★☆	無効ロジック、クレームチャート、専門家意見書の反証探索	証拠採否、訴訟戦略、privilege判断
技術評価・ライセンス	★★★★☆	技術主張、実験、数値、根拠のdue diligence	FTO・侵害・有効性の最終意見

先行技術調査

ここで最も重要なのは、**Deep ReviewとPatent Searchを分離すること**である。

Googleは「existing workとの比較」を機能として挙げるが、Deep ReviewにJ-PlatPat、USPTO、EPO、WIPO等の網羅的・最新特許コーパスや、優先日、ファミリー、法的状態、IPC/CPC/FI/F-termを保証する特許検索基盤があるとは公表していない。[16]

したがって、推奨パターンは、

J-PlatPat／Derwent／Patentfield／Amplified等でretrieval → Deep Reviewでreasoning／verification

である。J-PlatPatはINPITが世界の特許情報を無料で検索・閲覧できるサービスとして提供しており、Amplifiedも1.4億件超の特許を対象とする検索・ランキングを提供している。[17]

特に有望なのは、検索後に以下を問う使い方である。

> 「独立請求項1の各構成要件A〜Fについて、引用文献X/Y/Zのどの記載が対応するか。対応が弱い箇所、推論で補っている箇所、文献間結合が必要な箇所を区別して指摘せよ」

その結果を人間が原文献で再検証する。

Amplified自身の2026年のベンチマーク分析も、一般LLMは「特許をreasoningする用途」には向くが、「検索コンポーネント」そのものとして使うとrecall ceilingが低いと論じており、専用retrievalとLLM reasoningを組み合わせる方向性は合理的である。ただし、この比較はAmplifiedによる自社評価であり、第三者ベンチマークではない点を割り引く必要がある。 [18]

出願書類作成

もっとも即効性が高いユースケースである。

GoogleはDeep Reviewが数式、論理的整合性、セクション横断の依存関係を確認し、PATでは絶対値記号の欠落、不等号方向、notationの重複、off-by-oneなど細かな技術ミスまで捕捉したと報告している。 [15]

特許明細書で置き換えると、以下が対象になる。

チェック	Deep Reviewへの適合
独立請求項と従属請求項の論理関係	高
同一用語の表記・定義揺れ	高
請求項の要素と明細書の説明との整合	高
実施例の数値、単位、式、表の矛盾	高
図面番号と文章の食い違い	中。PDF parsing性能に依存
補正後クレームと元記載の対応	高。ただしmulti-file制約あり
サポート要件・実施可能要件	「候補問題の抽出」は高、法律判断は人間
新規性・進歩性	特許DBを併用しなければ不十分

特に「提出直前QA」として、従来の形式チェックとは異なる**意味レベルのlint**として使うのがよい。

審査対応

拒絶理由通知への応答では、

1. 審査官の主張、
2. 引用文献の記載、
3. 本願クレーム、
4. 本願明細書、
5. 補正案、
6. 意見書の論理、

の整合を確認する必要があり、本来Deep Reviewのmulti-document reasoningと非常に相性がよい。

ただし、2026年10月16日時点ではmulti-fileは「coming soon」である。 [19]

現時点での暫定策は、一つの統合テキスト/PDFへ

```
=== Document A: Current Claims ===
...
=== Document B: Specification excerpts ===
...
=== Document C: Office Action ===
...
=== Document D: Cited reference X ===
...
=== Draft Response ===
...
```

のように明示的な境界を付けてまとめる方法である。ただし元文書のページ・段落番号を維持し、Deep Reviewの指摘後に必ず元公報へ戻って確認する必要がある。

無効審判・訴訟準備

価値が大きい反面、**最も慎重な情報管理が必要な用途**である。

有用な役割は「red team reviewer」である。

例えば、自社の無効主張を入力し、

> 「相手方代理人としてこの無効論理を反証せよ。各反証について、前提誤り・引用文献の不足・claim construction依存・証拠不足・論理飛躍に分類せよ」

とすると、提出前の弱点探索に利用できる。

反対に、AI出力をそのまま「証拠」と扱うべきではない。AI指摘は**探索・検討のトリガー**であり、最終的な根拠は特許公報、論文、実験データ、技術常識資料、宣誓書・鑑定書等の原証拠で構成する必要がある。GoogleのPAT論文自身、AIレビューは人間の最終レビューを代替するものではなく、AIが誤った批判を作るため「原文書を検証する作業がAIレビューを検証する作業へ移る」危険を指摘している。 [20]

技術評価・ライセンス交渉

技術DDでは非常に使いやすい。

たとえばライセンサーの技術説明書、white paper、PoC結果について、

- 「性能30%改善」の母数は何か
- 比較条件は同じか
- 実験表と本文に矛盾はないか
- claimされた効果を全実施形態が実現するか
- 未検証の仮定を事実のように扱っていないか
- 技術的説明と契約上の表現にずれがないか

を洗い出す用途である。

Google自身、法律・金融文書について内部整合性、表との矛盾、曖昧な用語、downstream consequence、evidence groundingを利用例としているため、この用途は製品の設計思想にも近い。 [10]

期待効果と知財実務での評価方法

公開データから期待できるもの

PATのSTOC／ICMLパイロットで「明瞭性が改善した」と回答した利用者が85〜87%、90%以上がレビューをvery/mostly helpfulと評価したことから、**文書品質の第二査読としての有用性には相当の実績がある**。またICML回答者では35.4%が重大な理論上のgapを指摘されたと回答し、31%がPATの指摘を受けて新しい実験を行った。 [14]

ただしこれは自己申告アンケートであり、知財業務の「工数30%削減」「見逃し率50%減」といった効果を直接証明するものではない。

したがって導入時には、Googleの数字をROIに直接転用せず、自所データで測定すべきである。

推奨KPI

KPI	算出方法	PoCでの推奨目標例
Validated Precision	人間が有効と認定したAI指摘÷全AI指摘	>70%
Critical Recall	既知の重大問題のうちAIが検出した割合	>70%から開始
False-positive burden	誤指摘の確認に要した時間	<総削減時間の30%
Review time reduction	人手のみとの比較	20〜40%削減
Escape rate	最終版後に発見された問題率	ベースラインより20%低下
Citation grounding rate	根拠を原資料まで確認できたAI指摘比率	>90%
Human override rate	AI結論を専門家が覆した割合	継続監視
Major finding yield	1案件当たり重大な有効指摘数	業務別に測定
Confidentiality incidents	誤投入・不正共有	**0件**

ここで70%や20〜40%はDeep Reviewの公表性能ではなく、**導入可否判定のための本報告書の推奨管理値**である。

発見率を正しく測る

本番文書だけでは「本当は何件のミスがあったか」が分からないためrecallを測れない。

そこで、過去案件30〜50件程度を用いて、

- 実際に後から見つかった誤記・矛盾
- OAで問題となった論点
- クライアント／外国代理人から指摘された問題
- 意図的に埋め込んだsynthetic error

をgold set化する。

特許用SPOTを自所で作る発想である。

人間レビューのみをA群、Deep Review併用をB群としてblind比較すれば、工数と品質を同時に測れる。

導入リスク・法的留意点

機密情報と新規性

ここは「Googleは学習しないから安全」で終了してはいけない。

Deep Review UIはアップロード文書を学習、fine-tuning、モデル改善には使わないが、**UI表示用に1年間保存する**。 [16]

日本弁理士会「弁理士業務 AI 利活用ガイドライン」は、外部事業者の生成AIへ秘密情報を入力する行為について、第三者への秘密情報開示となり、弁理士法30条の守秘義務やNDAに抵触するおそれがあると指摘する。また、秘密保持規約のない第三者プラットフォームへの掲載は新規性喪失につながる可能性があるとして、サービス条件を確認するよう求めている。 [21]

重要なのは、**1年間保存＝直ちに公衆への公開＝直ちに新規性喪失、という意味ではない**ことである。実際の法的評価はアクセス権、秘密保持契約、サービス利用規約、開示先との関係、各国法等による。しかし、未公開発明を第三者クラウドへ預ける以上、事前のリーガル・セキュリティ審査は必要である。

特にJPAAは、クライアント秘密を生成AIに入力する場合について、**学習利用されない場合であってもクライアント側がAI利用を制限している可能性があるため、クライアント合意を取るべき**としている。 [21]

UI版とAPI版でデータ条件を分ける

これは実務上かなり重要である。

利用経路	確認できた条件
Deep Review UI	文書を学習・fine-tuning・モデル改善に使わない。UI表示目的で1年保存
Gemini API Unpaid Services	入出力を製品改善等に使用し得る。human reviewerが処理する可能性。秘密・機密・個人情報を入れないようGoogleが明記
Gemini API Paid Services	prompts/responsesを製品改善に使用しない。一定期間、安全・不正利用対策等のためlogging
Deep Review API固有条件	**未指定。契約時確認必須**

Gemini API Additional Termsは、unpaid quotaについて「Do not submit sensitive, confidential, or personal information」と明確にしている。一方、課金済みCloud Project経由のPaid Servicesでは入力・出力をGoogle製品改善に使わないとしている。 [5]

したがって、機密案件の業務利用では最低でも、

「無料APIだから試す」ではなく、Paid Servicesとしての契約状態、Deep Review固有条件、DPA、保持期間、アクセス権、リージョン等を確認してから利用

すべきである。

個人情報

個人情報保護委員会は、事業者が個人情報を生成AIへ入力するとき、特定した利用目的の必要範囲内であることを確認し、本人同意なしの個人データが応答生成以外に利用される可能性がある場合は、サービス事業者が機械学習に使わないこと等を確認するよう求めている。また利用規約とprivacy policyの確認も求めている。 [22]

特許業務では発明者氏名、担当者名、メール、署名、研究者情報、訴訟関係者情報などが混入しやすいため、入力前のpseudonymizationを標準化した方がよい。

AIの誤検出と責任

JPAAガイドラインは生成AIのhallucinationを明示し、生成情報の真偽確認を**必ず弁理士が行い、最終的に責任をもって検討・確認・判断する必要がある**としている。 [21]

Google PATでも、「正しいproofを誤りだと判断する」false positiveが報告されている。 [15]

したがってDeep Review出力には、

Confirmed / Plausible / Rejected / Needs expert review

の4状態を設け、AI自身の「critical」というラベルをそのまま案件重要度として扱わない運用が望ましい。

証拠性・説明可能性

無効審判・訴訟ではDeep Reviewの出力を「原証拠」とみなすべきではない。

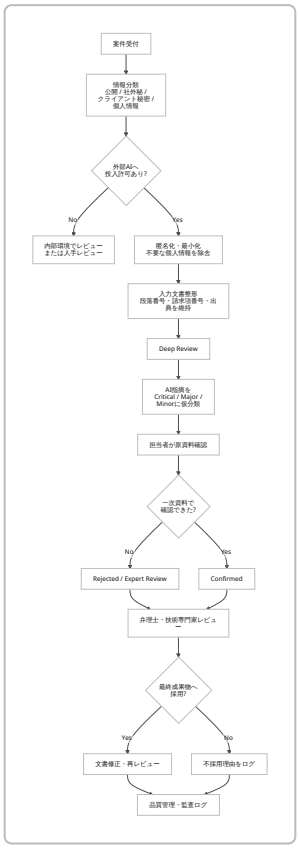
実務上は次を保存するとよい。

保存物	目的
入力文書のhash	どの版をAIがレビューしたか固定
文書version / date	証跡
Deep Review実行日時	再現性
利用モデル／サービスversion	後日の性能変更に備える
出力原本	監査
人間のaccept/reject理由	最終判断者の記録
根拠となった公報・論文	AIではなく一次証拠を残す
最終成果物との差分	AI寄与範囲の確認

モデルが更新されれば同じ入力で同じ結果になる保証はないため、**「再実行可能性」より「当時何を入力し、何が出て、誰がどう判断したか」を保存する**方が重要である。

運用設計・実装オプションと既存ツール比較

推奨運用フロー



入力データ整備

Deep Reviewの性能以上に、入力品質が実務成否を左右する。

望ましい形式は、

```
[DOCUMENT METADATA]
Matter: anonymized-2026-001
Purpose: Pre-filing technical consistency review
Priority date: [必要な場合のみ]
Jurisdiction: JP
Review scope: technical/logical consistency only

[CLAIMS]
Claim 1 ...
Claim 2 ...

[SPECIFICATION]
[0001] ...
[0002] ...
```

```
[EXAMPLES]
Example 1 ...

[KNOWN REFERENCES]
R1: JPxxxx...
R2: USxxxx...

[REVIEW QUESTIONS]
1. Claim terminology inconsistent with specification?
2. Any unsupported technical dependency?
3. Numerical contradictions?
4. Any statement not supported by supplied evidence?
5. Identify only issues for which exact source locations can be stated.
```

特に、モデルが勝手に法的判断を拡大しないよう「technical/logical consistency only」「根拠のページ／段落を必須」とするのが有効である。

ヒューマンインザループ

GoogleのPAT研究自体も、AIをauthor/reviewerの**補助ツール**として使い、最終責任を人間に残すRole 1／Role 2を区別している。Role 2ではAIのhallucinated critiqueを人間が検証する負担が生じるため、導入効果は「AIが何件指摘したか」ではなく「有効指摘を得るための総時間」で測らなければならない。 [20]

推奨する二段階レビューは、

Deep Review → 技術担当者の事実確認 → 弁理士／弁護士の法的判断

である。

重大案件なら、Deep Reviewの出力を担当者に先に見せず、人間レビューを独立して行ってから結果を照合するとconfirmation biasも抑えられる。

実装オプション

方式	構成	長所	短所	推奨用途
SaaS/UI	Deep Review Portal	即日利用、週2回無料	1年保存、手作業、統合性低い	公開資料PoC
API	DMS/案件管理→Deep Review API	自動化、ログ、匿名化を組込可能	API詳細・価格・SLA未指定	本格導入候補
オンプレ	自社LLM＋独自multi-agent	データ統制最大	**Deep Reviewそのものではない**、開発コスト大	極秘案件
ハイブリッド	社内前処理＋外部Deep Review＋社内検証	機密と性能のバランス	設計・運用が複雑	**最も推奨**

Deep Reviewそのもののオンプレミス版は、確認したGoogle公式情報では公表されていない。 したがって「オンプレDeep Review」という製品選択肢があるものとして計画してはいけない。 [19]

ハイブリッドの場合、未公開発明全体を送らず、

1. 社内で固有名詞・顧客名を置換、
2. レビューに不要なデータを除去、
3. 必要部分のみ外部検証、
4. 社内で元文書へ再マッピング、

とすることでリスクを下げられる。

既存ツールとの比較

Deep Reviewと特許検索製品は競合というより補完関係である。

ツール	主目的	特許検索	長文論理矛盾レビュー	AI/意味検索	価格の公開状況	公開精度	導入難易度
Google Deep Review	文書検証	△	◎	◎	週2レビュー無料、API固有価格未指定	PAT限定ベンチ89.7%	UI: 低 / API: 中
J-PlatPat	公的特許検索	◎	×	限定的	**無料**	該当なし	低
Patentfield / AIR	特許検索・分析・生成AI	◎	△	◎	BASIC 1万円/月、AIR +3万円/月〜/ID等	本調査で比較可能な統一指標なし	中
Amplified	AI prior-art search・IP knowledge	◎	△	◎	年払いで$500/月〜	ベンダー独自評価あり、DRとの直接比較不可	中
Derwent Patent Search	商用特許検索・分析	◎	△	◎	要見積	統一DR型精度指標なし	中〜高

J-PlatPatについてINPITは、日本だけでなく欧米等を含む世界の特許情報を無料で検索・閲覧できると説明している。 [23]

Patentfieldは2026年時点の公開料金としてFree、BASIC年10万円／月払い1万円、生成AI調査・分析オプションPatentfield AIRを月額3万円〜/ID、専用クラウドを月5万円〜・初期15万円〜などと掲載している。 [24]

Amplifiedは140 million超の特許を毎週更新し、AIランキング、boolean/filter、legal status等を提供し、価格は年払いで月500米ドルからとしている。 [25]

ClarivateのDerwent Patent Searchは、AI検索と専門家により整備された特許データを提供する商用検索プラットフォームで、2024年のAI検索発表では1億6千万件以上のpatent recordsを対象としていた。公開固定価格ではなく問い合わせ型である。 [26]

ここから導かれる実務上の最適構成は、

```
特許Retrieval
J-PlatPat / Patentfield / Amplified / Derwent
            ↓
候補文献・claim chart
            ↓
Deep Review
「論理・根拠・整合性」の検証
            ↓
弁理士 / 技術担当者
法律・技術の最終判断
```

である。

導入コスト、ROIとロードマップ

中規模特許事務所モデル

地域・人件費・案件構成が指定されていないため、以下を**試算上の仮定**とする。

中規模事務所として弁理士・技術者等のfee earner 15名前後、補助スタッフを含む総勢20名前後を想定する。Deep Review対象は出願、OA、無効／調査、技術評価等から月20〜100件。内部工数原価は8,000〜10,000円/時と仮置きする。

重要なのは、**Deep Review API固有価格は2026年10月16日時点の公式発表では未指定**である点である。Googleの一般Gemini APIにはtoken料金表が存在するが、Deep Reviewが1レビューあたり何token・何回のagent call・何回のgroundingを消費するかが公開されていないため、一般Gemini料金から正確なDeep Review単価を算出することはできない。 [27]

そこでAPI費を仮定して感度分析する。

ROI試算

シナリオ	月間レビュー	1件削減時間	内部時間単価	工数価値/月	仮定AI単価	AI費/月	運用QA費/月	純便益/月	初期費	単純回収
小規模PoC	20	0.5h	¥8,000	¥80,000	¥500	¥10,000	¥50,000	¥20,000	¥600,000	約30か月

シナリオ	月間レビュー	1件削減時間	内部時間単価	工数価値/月	仮定AI単価	AI費/月	運用QA費/月	純便益/月	初期費	単純回収
標準運用	**60**	**0.75h**	**¥8,000**	**¥360,000**	**¥2,000**	**¥120,000**	**¥100,000**	**¥140,000**	**¥1,200,000**	**約8.6か月**
高負荷	100	1.0h	¥10,000	¥1,000,000	¥5,000	¥500,000	¥200,000	¥300,000	¥2,000,000	約6.7か月

これは**価格予測ではなくwhat-ifモデル**である。

式は、

$$\text{月間便益} = \text{案件数} \times \text{一案件当たり削減時間} \times \text{内部時間単価}$$

$$\text{月間純便益} = \text{月間便益} - DeepReview\text{利用費} - \text{人間}QA\text{運用費}$$

である。

このモデルには「誤記を提出前に発見したことでOAや補正を回避した価値」「クライアントクレームの回避」「無効資料の早期発見」などは入れていない。そのため、品質事故回避に高い価値を置く事務所では実際の便益は上振れし得る。一方、false positive確認が多ければ削減時間は下振れする。

無料枠の経済効果と限界

10名が各自正当なユーザーとして週2回使う場合、理論上は月平均約86.7レビューになる。

しかし、これは「中規模事務所なら無料で本番運用できる」という意味ではない。

第一に、UIは1年間文書を保存する。第二に、組織管理、SSO、監査、DLP、SLAなどのenterprise条件は今回の公式発表では明確でない。第三に、未公開発明を含むクライアント秘密は、価格ではなく守秘条件によって利用可否を判断しなければならない。[4]

したがって無料枠は、**公開特許・公開論文・匿名化済み過去案件によるPoC枠**と考えるのが安全である。

推奨導入ステップ

短期：公開情報だけで性能を測る

期間目安は0〜3か月。

過去案件30〜50件を選定し、すでに公開済みの明細書、OA、引用文献等でgold setを作る。Deep Review UIの無料枠を使い、技術分野別にvalidated precision、critical recall、工数、false-positive確認時間を計測する。

この期間は未公開発明を投入しない。

中期：承認済み案件をAPI／ハイブリッド化

期間目安は3〜9か月。

Deep Review APIの正式仕様、料金、DPA、retention、region、security、利用ログ等を確認し、案件管理システムやDMSの前段に情報分類・匿名化処理を置く。

この段階では「出願前QA」「OA response QA」を主対象にする。

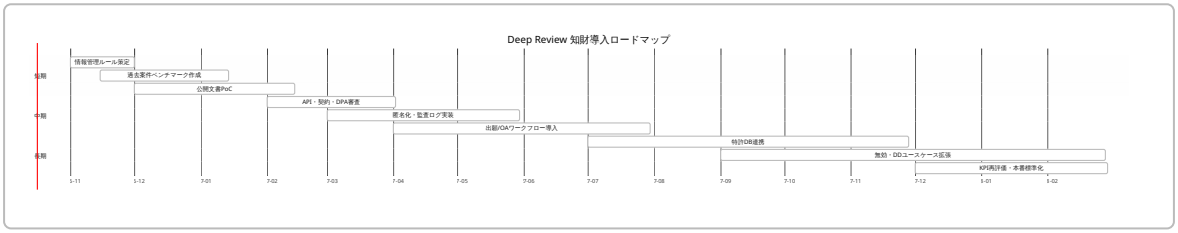
長期：特許検索システムと統合する

期間目安は9〜18か月。

特許DBから候補文献を取得し、

Retrieval → Claim mapping → Deep Review → Human verification

までをworkflow化する。



実務チェックリスト、想定Q&A、参考文献

運用開始前チェックリスト

情報管理

- [ ] 入力対象を「公開／社内限定／クライアント秘密／営業秘密／個人情報」で分類した
- [ ] 未公開発明を無料UIへ投入しないルールを定めた
- [ ] クライアント契約・NDAの生成AI条項を確認した
- [ ] 必要なクライアント同意を取得するフローを作った
- [ ] UIの1年間保存を情報セキュリティ部門が承認した
- [ ] API利用時はPaid/Unpaid Servicesのどちらになるか確認した
- [ ] APIの保持期間・DPA・data location・subprocessor条件を確認した
- [ ] 個人情報を削除・仮名化する処理を設けた

品質管理

- [ ] Deep Reviewの指摘を最終判断として利用しない
- [ ] 指摘ごとに元文書のページ／段落／請求項番号を確認する
- [ ] AIが挙げる特許・論文の実在を一次DBで確認する
- [ ] 出願日、公開日、優先日、法的状態をAI回答から直接採用しない
- [ ] validated precisionとfalse positive率を月次計測する

- [ ] 技術分野別の精度差を測定する
- [ ] モデル／サービス更新後にgold setを再実行する

証跡・ガバナンス

- [ ] 入力ファイルhashを残す
- [ ] AI実行日時を残す
- [ ] モデル／サービスversionを可能な範囲で残す
- [ ] AI出力原文を保存する
- [ ] 人間が採用／棄却した理由を記録する
- [ ] 最終責任者を案件ごとに指定する
- [ ] AI利用可否を案件管理システム上で識別できるようにする

想定Q&A

Q. Deep Reviewは特許調査ツールを置き換えられるか。

A. 置き換えない。Deep Reviewはdocument verifierであり、J-PlatPatやDerwent、Patentfield、Amplifiedのような特許retrieval基盤とは役割が異なる。最適なのは検索ツールで集めた文献をDeep Reviewにreasoningさせる構成である。Googleはexisting work比較機能を説明しているが、網羅的な特許コーパス、優先日・legal status、IPC/CPC等の検索品質を保証していない。 28

Q. 「89.7%精度」なら人間レビューを減らしてよいか。

A. その解釈はできない。89.7%はMath/CSの式・証明エラー29件に限定したPAT評価で、特許案件ではない。さらにoriginal SPOTとはgraderも違う。Google自身、誤って正しい議論を否定するfailureを報告している。 29

Q. 未公開発明をDeep Reviewへそのまま入れてよいか。

A. 現段階では推奨しない。UIは学習利用しないが、文書を1年間保存する。JPAAはクライアント秘密を外部AIに入力する場合、学習の有無に関係なくクライアント合意を取るべきとしている。契約・セキュリティ審査後に判断すべきである。 4

Q. Googleが「学習に使わない」と言っているのに何が問題なのか。

A. 「モデル学習に使わない」と「第三者クラウドへ情報を保管してよい」は別問題である。守秘義務、NDA、個人情報、データ所在地、retention、アクセス権、クライアント社内規程等を別に確認する必要がある。JPAAもこの区別を実質的に要求している。 21

Q. 無料APIを業務利用すればよいのではないか。

A. 機密案件では避けるべきである。Gemini API一般規約ではUnpaid Servicesの入力・出力を製品改善等に利用し、人間がレビューする場合があり、秘密情報を投入しないよう明記されている。Paid Servicesでは製品改善に使わない。Deep Review APIの具体的契約条件は別途確認が必要である。 5

Q. Deep Review APIは今すぐ実装可能か。

A. GoogleはGemini API keyによるprogrammatic integrationを公式発表している。ただし、本調査で確認できた公式公開情報にはDeep Review専用endpoint、request/response schema、rate limit、専用価格、SLA

の詳細がない。そのため2026年10月16日時点ではこれらを**「未指定」**として要件定義すべきである。なお、Googleの「Deep Research agent API」と「Deep Review API」は名前が似ているが別機能であり、混同してはいけない。 30

Q. API単価が不明なのにROIを計算できるのか。

A. 一点予測はできない。そのため本報告では500円、2,000円、5,000円/レビューという仮定を置いた感度分析としている。正式導入時は見積価格に置換すれば同じモデルで再計算できる。

Q. 訴訟でDeep Reviewの出力を証拠として使えるか。

A. AI出力自体を技術的事実の原証拠とする設計は推奨しない。AIは論点探索・反証候補生成に使い、最終主張は原特許、刊行物、実験、専門家意見等で裏付ける。手続上の証拠能力・privilege/work-product等は法域と案件ごとに弁護士が判断すべきである。

Q. オンプレ版を待つべきか。

A. GoogleはDeep Reviewのオンプレ版を現時点で発表していない。極秘案件で閉域性が必須なら、自社LLM／private modelによる別のmulti-agent reviewを構築するか、公開情報・匿名化部分のみDeep Reviewへ渡すハイブリッド構成が現実的である。 19

Q. 最初に何をレビューするのがよいか。

A. 最優先は、すでに公開済みで「後からどんな問題が判明したか」が分かっている過去出願・OA案件である。これなら秘密性を低く保ちつつ、recallとprecisionを実測できる。

総合提言

Deep Reviewは、知財業務において「特許版の自動査読者」としてかなり有望である。特に従来の生成AIが得意とした要約・文章生成ではなく、**人間の成果物を敵対的に読み、矛盾・根拠不足・技術的な穴を探す**方向へAI利用を移せる点が重要である。

これは知財業務と相性がよい。特許実務では、「綺麗な文章を生成する能力」より「一見正しく見える文書に潜む小さな不整合を見つける能力」の方が価値を持つ場面が多いためである。

ただし、現時点でDeep Reviewを**自律的な特許判断AI**と位置付けるのは早い。特許固有の精度ベンチマークはなく、multi-fileもまだ未実装で、API仕様・専用料金・SLAも明らかでない。UIは文書を1年間保存し、APIではPaid/Unpaidのデータ取扱条件にも差がある。 31

したがって、中規模特許事務所に対する推奨アーキテクチャは、

> **特許専用検索基盤 + Deep Reviewによる第二査読 + 弁理士による最終判断**

である。

短期的には公開過去案件でPoCを実施し、30〜50案件のgold setで独自のprecision/recallと工数を測る。中期的にはPaid APIの契約条件が確認できた段階で匿名化・監査ログを組み込んだハイブリッド運用へ進み、長期的には特許検索結果、クレームチャート、OA、明細書を連携させる。この順序であれば、生成AIの最大の

弱点である「もっともらしい誤り」と、知財業務最大のリスクである「未公開情報の漏洩」の双方を管理しながら、Deep Reviewの強みを利用できる。

主要参考文献・一次情報URL

指定・主要情報源

1. Google Research, Analyzing technical documents with Deep Review, 2026-10-14
   https://research.google/blog/analyzing-technical-documents-with-deep-review/ [6]

2. Google Research, Deep Review Portal
   https://research.google/deepreview/

3. CNET Japan, 「論文の矛盾を見つけるAI『Deep Review』、Googleが公開　週2回まで無料」, 2026-10-16
   https://japan.cnet.com/article/35253606/
   同記事の配信版でタイトル・日付・概要を確認。 [32]

4. Jayaram et al., Towards Automating Scientific Review with Google's Paper Assistant Tool, arXiv: 2606.28277
   https://arxiv.org/abs/2606.28277
   HTML: https://arxiv.org/html/2606.28277v1 [33]

技術・関連研究

1. Son et al., When AI Co-Scientists Fail: SPOT—a Benchmark for Automated Verification of Scientific Papers, 2025
   https://arxiv.org/abs/2505.11855 [34]

2. de Marneffe, Rafferty & Manning, Finding Contradictions in Text, ACL 2008
   https://aclanthology.org/P08-1118/ [35]

3. Williams et al., A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference (MultiNLI), 2018
   https://aclanthology.org/N18-1101/ [36]

4. Zhu et al., DeepReview: Improving LLM-based Paper Review with Human-like Deep Thinking Process, ACL 2025
   ※Google Deep Reviewとは別プロジェクト。
   https://aclanthology.org/2025.acl-long.1420/ [37]

Google API・データ取扱い

1. Google, Gemini API Additional Terms of Service
   https://ai.google.dev/gemini-api/terms [5]

2. Google, Gemini API Key Documentation
   https://ai.google.dev/gemini-api/docs/api-key

3. Google, Gemini Developer API Pricing
   ※Deep Review専用料金ではない。
   https://ai.google.dev/gemini-api/docs/pricing [38]

4. Google, Gemini API Document Understanding
   https://ai.google.dev/gemini-api/docs/document-processing [39]

日本の知財・法務ガイド

1. 日本弁理士会, 「弁理士業務 AI 利活用ガイドライン」, 2025年4月
   https://www.jpaa.or.jp/cms/wp-content/uploads/2025/04/AIservices-guideline.pdf
   守秘義務、NDA、新規性、個人情報、クライアント同意、hallucinationと人間の最終責任について特に重要。 [21]

2. 個人情報保護委員会, 「生成AIサービスの利用に関する注意喚起等について」
   https://www.ppc.go.jp/files/pdf/230602_kouhou_houdou.pdf [22]

3. INPIT, 特許情報プラットフォーム J-PlatPat
   https://www.j-platpat.inpit.go.jp/
   案内: https://www.inpit.go.jp/j-platpat_info/index.html [40]

比較対象ツール

1. Patentfield 料金プラン
   https://patentfield.com/pricing [41]

2. Amplified AI
   https://www.amplified.ai/
   Pricing: https://www.amplified.ai/pricing/ [25]

3. Clarivate, Derwent Patent Search
   https://clarivate.com/intellectual-property/derwent/patent-search/ [42]

調査基準日：2026年10月16日（日本時間）。 Deep Reviewは公開から2日しか経っておらず、特にAPI仕様、価格、SLA、データ保持条件、multi-file対応は今後短期間で変更される可能性が高い。日本弁理士会も生成AIの利用規約・プラン仕様は頻繁に変更され得るため最新条件を継続確認するよう求めており、本番導入時には上記の「未指定」項目を契約・公式ドキュメントで再確認することが必須である。 [21]

---

[1] [32] 論文の矛盾を見つけるAI「Deep Review」、Googleが公開　週2回まで無料 | チバテレ＋プラス
https://www.chiba-tv.com/plus/detail/2026102119145

[2] [4] [6] [8] [10] [11] [16] [19] [27] [28] [30] [31] Analyzing technical documents with Deep Review
https://research.google/blog/analyzing-technical-documents-with-deep-review/

[3] [12] [14] [15] [20] [29] [33] Towards Automating Scientific Review with Google’s Paper Assistant Tool
https://arxiv.org/html/2606.28277v1

[5] Gemini API Additional Terms of Service | Google AI for Developers
https://ai.google.dev/gemini-api/terms

7 論文の矛盾を見つけるAI「Deep Review」、Googleが公開 週2回まで無料 | IT NEWS
https://news.minory.org/707576.html

9 35 Finding Contradictions in Text - ACL Anthology
https://aclanthology.org/P08-1118/?utm_source=chatgpt.com

13 37 DeepReview: Improving LLM-based Paper Review with Human-like Deep Thinking Process - ACL Anthology
https://aclanthology.org/2025.acl-long.1420/

17 23 ［INPIT］特許情報プラットフォーム（J-PlatPat）
https://www.inpit.go.jp/j-platpat_info/index.html?utm_source=chatgpt.com

18 Benchmarking AI search components and LLMs
https://www.amplified.ai/en/blog/07042026/benchmarking-components-llms/?utm_source=chatgpt.com

21 jpaa.or.jp
https://www.jpaa.or.jp/cms/wp-content/uploads/2025/04/AIservices-guideline.pdf

22 生成AIサービスの利用に関する注意喚起等について -個人情報保護委員会-
https://www.ppc.go.jp/files/pdf/230602_kouhou_houdou.pdf

24 41 料金プラン
https://patentfield.com/pricing?utm_source=chatgpt.com

25 Amplified - Home
https://www.amplified.ai/

26 42 Derwent Patent Search (formerly Derwent Innovation)
https://clarivate.com/intellectual-property/derwent/patent-search/?utm_source=chatgpt.com

34 When AI Co-Scientists Fail: Spot—a Benchmark for ...
https://arxiv.org/html/2505.11855v1?utm_source=chatgpt.com

36 A Broad-Coverage Challenge Corpus for Sentence ...
https://aclanthology.org/anthology-files/pdf/N/N18/N18-1101.pdf?utm_source=chatgpt.com

38 Gemini Developer API pricing
https://ai.google.dev/gemini-api/docs/pricing?utm_source=chatgpt.com

39 Document understanding - Interactions API
https://ai.google.dev/gemini-api/docs/document-processing?utm_source=chatgpt.com

40 特許情報プラットフォーム｜J-PlatPat [JPP]
https://www.j-platpat.inpit.go.jp/?utm_source=chatgpt.com