Google「Deep Review」調査報告

論文の矛盾を見つける AI の事実確認と、知財業務での活用方法・注意点

2026 年 10 月 16 日

Claude Opus 5.5

本報告は、Google が公開した文書検証 AI「Deep Review」を報じた CNET Japan の記事[1]を起点に、公式発表などの一次情報で事実関係を確認し、知財業務での活用方法と注意点を整理したものである。一次情報等で確認した事実と筆者の分析・推論を区別し、後者には【分析】または【推測】と付した。

要旨

- Deep Review は、Google Research が 2026 年 10 月 14 日に一般公開した、技術文書の論理・数式・根拠の不整合を指摘する AI エージェントシステムである[2]。全ユーザーに週 2 回の無料レビューがあり、それ以上は自分の Gemini API キーを接続して使う[2]。
- 前身は学会投稿前の論文チェック用「Paper Assistant Tool（PAT）」で、STOC 2026、ICML 2026、NeurIPS 2026、ICLR 2027 の 4 学会で 1 万 6000 件超の原稿に提供された[2]。
- 公表されている精度データはすべて PAT（旧システム）のもので、公開版そのものの性能評価は示されていない[3,4]。
- 【分析】知財業務では、公開済みの文献（他社特許、引用文献、学術論文）の矛盾の洗い出しにはすぐ使える。一方、未公開の発明や自社明細書のドラフトは、Deep Review 専用の秘密保持条件が確認できず、データの扱いについて公式情報の間に表現の差もあるため、現時点では投入を控えるべきである。

1. Deep Review の概要（確認済みの事実）

1.1 基本情報

項目	内容
公開日	2026 年 10 月 14 日。米マウンテンビューで開いた Google のイベント「Research@」で発表[2,5]
提供主体	Google Research。Google Research、Google DeepMind、Google Cloud の共同開発[2]
利用方法	research.google/deepreview に Google アカウントでアクセス[2]
対象地域	全世界。日本でも利用可能と報道[2,6]
言語	日本語など英語以外への対応は、公式情報で確認できず
料金	週 2 回無料。超過分は自分の Gemini API キーを接続し、Gemini API の標準料金のみを払う（Deep Review 自体の追加料金なし）。FAQ の目安は 40 ページの文書で約 3 ドル（Low、Gemini 3.8 Flash）～約 20 ドル（Extra High、Gemini 4 Argon）[4,7]
無料枠のモデル	Gemini 4 Argon（計算量 Medium）と報道[5]
入力形式	公式ブログは「当初は PDF とプレーンテキスト、複数ファイル入力は近日対応」とする[2]。サービスページ（10 月 15 日時点）は PDF・テキスト・Google ドキュメント・Google スライドを挙げ[4]、R&D World は Word・PowerPoint・Markdown も挙げる[5]。情報源の間で差がある
出力	主要な問題・軽微な問題・改善提案・要約を、本文の該当箇所に紐づけて返す。所要時間は約 10～20 分（開発者のイベントでの発言）[5]
API	Deep Review API あり[2]
関連	開発環境「Antigravity」に、Deep Review に着想を得た文書レビュー手法を搭載（/teamwork-preview コマンド）[2]

1.2 仕組み

公式ブログによれば、Deep Review は次の 4 段階で文書を検証する[2]。

①**文書分割**：分割担当のエージェントが、序論・方法・実験などの論理単位に文書を分ける。区分は重複してよいが、全体を漏れなく覆う。

②**計算資源の配分**：情報密度に応じて計算量を割り振る。数学的証明が密な部分には、序論や関連研究より多く割り当てる。

③**並列の深い検証**：各エージェントは担当箇所に集中しつつ、文書全体を参照できる。このため章をまたぐ矛盾も追える。

④**統合と外部照合**：統合エージェントが各区分の報告をまとめ、章間の整合性を最終確認する。

外部検索などで照合し、ハルシネーションを防ぐ。

検出の対象は、論理の一貫性と数学的記述の検証、改善提案、事実の裏付け、既存研究との比較である[2]。PAT のパイロットでは、証明の欠落、不等号の向きの誤り、記法の重複、off-by-one エラーなどが指摘された。機械学習の学会向けには、実験設計の交絡要因や比較・実験の不足も指摘できるよう拡張された[3]。論文以外の用途として、公式ブログは設計書、プレゼン資料、法務・財務文書（条項間・データ表間の整合性、曖昧な用語、主張の根拠照合）を挙げる[2]。

1.3　精度・評価データ

撤回や正誤表が出た論文を集めた SPOT ベンチマークから、数学・CS 分野の「数式／証明」の誤りを含む 26 本・29 件を抜き出して評価した結果、検出率は PAT（Gemini 3.1 Pro）が 89.7%、Gemini 3.1 Pro の単発応答が 55.2%、元の SPOT 論文の最高値が 21.1%だった[3]。ただし採点には、論理的に同じ誤りかを判定する独自の LLM 採点器（著者が全件を人手で監査）を使っており、元論文の数値とは直接比較できないと論文自身が断っている[3]。

パイロットに参加した著者へのアンケート結果は次のとおりである[3]。

質問	STOC（n=124）	ICML（n=733）
また使いたい	97%	92.1%
非常に／おおむね役立った	92.7%	90.7%
指摘がおおむね／すべて根拠あり	55.8%	64.8%
修正に 1 時間超を要する理論上の重大な欠陥を指摘された	11.6%	35.4%
新たに実験を行った	—	31%

PAT のパイロットは「Gemini 2.5 Deep Think」の高度版で動いていた[3]。公開版は Gemini 4 Argon に対応し、性能はさらに上がったと Google は説明している[2]。しかし、公開版そのものの評価データは公表されていない[4]。

1.4　既知の限界

PAT 論文は、パイロットで多く報告された課題として、①日付のハルシネーションと知識の古さ、②PDF の解析エラー、③推論の失敗によって正しい証明や議論を誤りと判定すること、の

3 点を挙げる。①②は検索と解析の改善で対処済みとし、③は LLM を使うシステムに共通する問題として改善を続けているという[3]。指摘が「おおむね根拠あり」とした回答が STOC で約56%にとどまる点からも、指摘の相当部分を人間がふるい分ける必要がある[3]。

1.5 データの取り扱い（公式情報の間に表現の差あり）

- **公式ブログ**：アップロードした文書は、Google のモデルの訓練・ファインチューニング・改善に一切使わない。保存は 1 年間で、目的は Deep Review の画面上での表示のみ[2]。
- **サービスページ**：研究チームが Deep Review の改善にデータを使うことを、利用者が許可できるオプトインの選択肢がある[4,7]。関連研究を探すために Web 検索を行う場合がある旨の記載も検索抜粋で見られるが、本文は未確認[7]。
- **確認できなかった点**：Deep Review 専用の利用規約・プライバシー通知、利用者自身による文書削除の可否、Web 検索時に文書の内容が検索語として送られるか、自分の API キーを使う場合にどの規約が適用されるか。
- **参考（Gemini API の規約）**：無料枠（Unpaid Services）では、入力内容を製品改善に使い、人間のレビュアーが読む場合があるとして、機密情報を送らないよう求めている。課金が有効な Cloud プロジェクト経由の有料サービスでは、プロンプトを製品改善に使わない[8]。
- **パイロット時（PAT）**：訓練用のデータは集めず、各プログラム終了後に論文を速やかに削除した[2,9]。公開版の「1 年保存」とは扱いが異なる。

1.6 類似ツールとの比較

ツール	特徴
Deep Review（Google）	多エージェント構成と推論スケーリング。論文以外の文書にも対応。週 2 回無料[2]
Refine（refine.ink）	経済学者 Benjamin Golub 氏らが 2025 年に共同創業[10]。推論・計算・参照の誤りを指摘する。2026 年 8 月、米国経済学会と計量経済学会が出版プロセスでの利用を発表[11]
Black Spatula Project	オープンソースのボランティア主体のプロジェクト。約 500 本を分析し、誤検出率は約 10%と報じられた[12]
YesNoError	Matt Schlicht 氏が創業し、独自の暗号資産で資金を調達。開始から 2 か月で 3 万 7000 本超を分析したと報じられた[12]

2. 評価・評判

- **肯定的な声**：パイロット参加者からは、何か月も見逃していたアルゴリズムのバグを見つけてもらったという声が寄せられた。カリフォルニア大学アーバイン校のヴィジェイ・ヴァジラニ特別教授も、アルゴリズムの致命的なバグを指摘されたと高く評価している[3]。NeurIPS プログラムチェアの Nika Haghtalab 氏（UC Berkeley）は、見た目だけ整って中身の弱い論文が作りやすくなったなかで、投稿前のフィードバックは査読者の時間を守ると述べた。さらに、詳細な検証を AI に任せれば、研究者は新規性や重要性の評価に時間を割けるとしている[5]。
- **懸念**：正しい証明を誤りとする誤検出がある[3,5]。PAT 論文自身も、体裁だけ強く見える論文の増加、査読者のスキル低下、AIレビューを狙った敵対的な操作、計算資源へのアクセス格差を課題に挙げる[3]。外部の研究では、AI 査読ツールは論文の書き換えで高い評価に誘導でき、AI が書いたレビューは互いに似通って多様性を欠くと報告されている[13]。
- **メディアの論調**：報道はおおむね事実中心である。Google が人間の専門家の判断を代替するものではなく、検証を補助するツールと位置づけている点を共通して伝えている[6,14]。

3. 知財業務での活用方法【分析】

3.1 前提となる事実

Deep Review は PDF とテキストを受け付け[2]、法務文書の条項間・データ表間の整合性チェックを公式の用途に掲げている[2]。したがって、特許公報の PDF を投入すること自体は技術的にできる可能性が高い。一方、特許明細書での評価事例、日本語への対応、特許法上の要件を判断する能力は、公式には示されていない。

3.2 すぐ使える用途（公開文書が対象で、守秘リスクが小さい）

(1) **無効化・FTO 調査での他社特許の弱点探し**：公開済みの特許公報について、請求項と実施例の不整合、数値範囲と実施例データの矛盾、用語の定義の揺れ、図面と本文の食い違い、効果の主張と実験結果の乖離を洗い出す。明確性要件（特許法 36 条 6 項 2 号）、サポート要件（同項 1 号）、実施可能要件（同条 4 項 1 号）を攻める手がかりになりうる。AI の指摘はあくまで「仮説リスト」であり、要件を満たすかの判断は実務家が行う。

(2) **先行技術としての学術論文の精査**：引用候補の論文の数式・数値・表の整合性やデータの信頼性を確かめる。外部照合の機能は、論文中の引用文献が実在するかを確かめる補助にもなる。主引例の開示内容に誤りがあれば、新規性・進歩性の議論への影響を前もって把握できる。

(3) **審査対応（拒絶理由の引用文献の読み込み）**：審査官が引いた文献の記載に内部矛盾がないか、実施例が主張どおりの開示になっているかを確かめ、反論の材料を探す。

(4) **鑑定・係争での技術文書レビュー**：公開済みの相手方特許、技術文献、標準規格文書などの長文について、章をまたぐ不整合を検出する。ただし、相手方から開示された秘密資料や保護命令の対象文書は、3.3 と同じ扱いにすべきである。

3.3　効果は大きいが、社内判断が先の用途（未公開情報が対象）

(5) **自社明細書のドラフトチェック**：請求項と実施例の対応、数値範囲の一貫性、用語の統一、符号の対応、計算式の誤りを確かめる。Deep Review の「長文・章間の整合性」という強みと最も相性がよいが、出願前のドラフトは未公開の発明そのものである。

(6) **研究者の論文発表前チェックと発明発掘**：論文の質を高めながら、知財部が出願すべき発明の漏れを確かめる運用は魅力的である。ただし、アップロードは発表（公開）より前の時点で起きる。

3.4　特許明細書に使えるか（事実と推測の区別）

論点	状態
PDF・テキストの受付	確認済み[2]
法務文書の整合性チェックが公式の用途	確認済み[2]
特許明細書での評価・事例	確認できず
日本語文書への対応	確認できず。【推測】Gemini の多言語能力から一定の処理はできると見込まれるが、検証データはない
特許法上の要件の判断	【推測】一般的な論理整合性の指摘にとどまる可能性が高い。サポート要件などの法的評価は人間が行う必要がある
複数文書の横断比較（自社請求項と引用文献の対比など）	現時点では不可。複数ファイル入力は近日対応予定[2]

4. 注意点

(1) **守秘・新規性喪失のリスク**：日本弁理士会のガイドラインは、入力した情報が学習に使われ第三者への出力に反映されるおそれがあるか否かは「新規性の喪失や秘密保持契約等の遵守に密接に関係する」とする[15]。そのうえで、秘密保持規約のない第三者のプラットフォームに情報を載せると新規性を失う可能性があると注意を促し、クライアントの秘密情報を入力する場合は学習の有無によらず合意を取るべきとしている[15]。Deep Review は訓練に使わないと表明している[2]が、専用の秘密保持条件は確認できず、オプトインの記載もある[4]。【分析】守秘義務を負う相手への開示は直ちに公知とはならないとの考え方もあり、評価は分かれうる。それでもリスクを否定しきれない以上、未公開の発明は投入しないのが安全側の判断である。新規性喪失の例外（特許法 30 条）は要件と手続が限られるため、頼るべきではない。

(2) **規約上のデータの扱い**：オプトインがオフになっているかを必ず確かめる。自分の API キーを使う場合、そのキーが無料枠か、課金が有効なプロジェクトかで適用条件が変わりうる[8]。文書は 1 年間保存され、削除の方法は確認できていない。

(3) **輸出管理・安全保障**：USPTO の AI ツール利用ガイダンス（2024 年 4 月）は、国外のサーバーを使う AI ツールが、国家安全保障・輸出管理・外国出願許可の問題を生じうると指摘している[16,17]。Gemini API の有料サービスでも、プロンプトは Google の施設がある国で一時的に保存されうる[8]。該非判定の対象となる技術情報の投入には注意が必要である。

(4) **無料枠の制約**：週 2 回では日常業務には足りない。本格的に使うには API の従量課金が必要で、目安は 40 ページあたり約 3〜20 ドルである[4]。

(5) **誤検出・見落とし**：最良の条件でも約 1 割の誤りを見落とし、正しい記載を誤りとする例もある[3]。「指摘がない＝問題がない」ではない。

(6) **人間による最終確認**：USPTO のガイダンスも日本弁理士会のガイドラインも、AI の生成物を確認する義務と最終責任は実務家にあるとしている[15,17]。

5. 推奨事項【分析】

(1) **公開文書から試す**：他社の特許公報、拒絶理由の引用文献、先行技術の論文で試行し、指摘の精度と日本語文書での挙動を社内で検証する。日本語公報と英語のファミリー公報を並べ

て比べると効果的である。

(2) **社内ルールを先に整える**：未公開の発明、ドラフト、研究原稿を投入してよいかを、知財部と情報セキュリティ部門で事前に決める。少なくとも、①Deep Review 専用の規約の有無、②オプトインの初期設定、③削除の手段、④API キー経由で使う場合の適用規約、を書面で確認するまでは投入を禁じる。

(3) **補助ツールと位置づける**：AI の指摘は「調査の仮説」として扱い、36 条の各要件や 29 条の新規性・進歩性といった法的評価は必ず実務家が行う運用にする。

(4) **複数ファイル対応を注視する**：複数ファイルの入力が実装されれば、自社の請求項と複数の引用文献を横断してチェックでき、知財実務での価値が大きく上がる[2]。

6. 今後の影響・示唆【分析】

- **投稿前チェックの標準化**：大手による無料提供は、論文の投稿前チェックを事実上の標準にする可能性がある。先行技術文献の記載の質は上がる一方、AI で体裁を整えた文献も混じるため、知財部が文献の信頼性を見極める重要性はむしろ高まる。
- **攻防の非対称性**：他社特許の内部矛盾を安く大量に洗い出せるようになると、無効化調査のコストが下がる。裏返せば、自社の明細書の品質がこれまで以上に問われる。後者を同じ AI で担保するには守秘の課題を解く必要があり、秘密保持契約付きの企業向け提供形態が出るかが鍵になる（本調査では、そうした発表は確認できなかった）。
- **特許庁・代理人への波及**：審査側・代理人側の AI 利用ルール（開示、守秘、輸出管理）をめぐる議論が、検証型の AI にも及ぶとみられる。

7. 本調査の限界

- サービスページは JavaScript で描画されるため本文を直接取得できず、FAQ の記載は検索結果の抜粋と NeoTeo の記事[4]で確認した。
- 起点記事（Yahoo!ニュース配信）は自動取得が拒否され、本文を直接確認できていない[1]。
- パイロットの件数は、公式ブログでは 1 万 6000 件超[2]、Google 公式 X の投稿（GIGAZINE が引用）と R&D World では 1 万 5000 件超[5,14]と食い違っている。

- 精度データはすべて旧システム（PAT）とコンピュータ科学分野のもので、特許・化学・生命科学など他分野での性能は不明である。
- 新規性喪失や守秘義務に関する記述は一般的な整理であり、個別案件への法的助言ではない。

参考文献

URL の最終確認日はいずれも 2026 年 10 月 16 日。※印は本文を直接確認できなかった文献を示す。

[1] 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」Yahoo!ニュース配信. https://news.yahoo.co.jp/articles/d19c127fffd7b11fe2bdeba4f86d68dd6c6c143a ※自動取得が拒否されたため本文・配信日時は未確認（検索結果の抜粋のみ確認）。

[2] Rajesh Jayaram, Vincent Cohen-Addad. "Analyzing technical documents with Deep Review." Google Research Blog, 2026 年 10 月 14 日. https://research.google/blog/analyzing-technical-documents-with-deep-review/

[3] Rajesh Jayaram, Drew Tyler, David Woodruff, Corinna Cortes, Yossi Matias, Vahab Mirrokni, Vincent Cohen-Addad. "Towards Automating Scientific Review with Google's Paper Assistant Tool." arXiv:2606.28277（プレプリント, 2026 年 6 月）. https://arxiv.org/abs/2606.28277 （v1 の HTML 版で内容を確認）

[4] NeoTeo. "Google announced worldwide public access to Deep Review." 2026 年 10 月 15 日. https://www.neoteo.com/en/google-announced-worldwide-public-access-to-deep-review

[5] Brian Buntz. "Google launches Deep Review AI tool to help researchers catch manuscript errors." R&D World, 2026 年 10 月 14 日. https://www.rdworldonline.com/google-launches-deep-review-ai-tool-to-help-researchers-catch-manuscript-errors/

[6] BigGo Finance. "Google Opens Personal Use of AI-Generated Apps, Launches Free AI Paper Verification Tool." 2026 年 10 月 15 日（UTC）. https://finance.biggo.com/news/7a9d9ad7-1c5b-46e6-860f-c320f7e0d412

[7] Google Research「Deep Review」サービスページ（FAQ を含む）. https://research.google/deepreview ※JavaScript で描画されるため本文を直接取得できず。記載内容は検索結果の抜粋および NeoTeo の記事で確認。

[8] Google. "Gemini API Additional Terms of Service."（2026 年 3 月 23 日発効、ページ最終更新 2026 年 4 月 28 日）. https://ai.google.dev/gemini-api/terms

[9] Rajesh Jayaram, Vincent Cohen-Addad, Alekh Agarwal, Miroslav Dudik, Sharon Li, Martin Jaggi. "ICML Experimental Program using Google's Paper Assistant Tool (PAT)." ICML Blog, 2026 年 1 月 14 日. https://blog.icml.cc/2026/01/14/icml-experimental-program-using-googles-paper-assistant-tool-pat/

[10] Wikipedia（英語版）「Benjamin Golub」. https://en.wikipedia.org/wiki/Benjamin_Golub ※二次情報。

[11] Marginal Revolution. "Solve for the (Refine) equilibrium." 2026 年 8 月 6 日（Refine による提携発表を引用）. https://marginalrevolution.com/marginalrevolution/2026/08/solve-for-the-refine-equilibrium.html

[12] Elizabeth Gibney. "AI tools are spotting errors in research papers: inside a growing movement." Nature, 2025 年. https://www.nature.com/articles/d41586-025-00648-5 ※原文は未確認。URL と記載内容は、同記事を紹介する複数の二次情報で確認。

[13] Ananya. "AI tools meant to vet science are surprisingly easy to fool." Science News, 2026 年 7 月 8 日. https://www.sciencenews.org/article/ai-tools-science-peer-review-problems

[14] GIGAZINE（英語版）. "Google has made 'Deep Review,' an AI-powered tool for verifying errors in academic papers and technical documents, publicly available for free use up to twice a week." 2026 年 10 月 15 日. https://gigazine.net/gsc_news/en/20261015-google-deep-review/

[15] 日本弁理士会「弁理士業務 AI 利活用ガイドライン」令和 7 年（2025 年）4 月. https://www.jpaa.or.jp/cms/wp-content/uploads/2025/04/AIservices-guideline.pdf

[16] United States Patent and Trademark Office. "Guidance on Use of Artificial Intelligence-Based Tools in Practice Before the United States Patent and Trademark Office." 89 Fed. Reg. 25609（2024 年 4 月 11 日）. ※原文は未確認。内容は下記 Crowell & Moring の解説で確認。

[17] Crowell & Moring. "USPTO Embraces AI Tools in New Guidance." Client Alert, 2024 年 5 月 3 日. https://www.crowell.com/en/insights/client-alerts/uspto-embraces-ai-tools-in-new-guidance