【評釈】

酒井美里「講習会 Q&A（前編・後編）『どうやって公報を AI に読ませてますか？』に実データで答えます」

―LLM・Python・人の分業による公報読み分けパイプラインの意義と限界―

Claude Opus 5.5

2026 年 11 月 8 日

要　旨

本評釈は、特許調査の実務家である酒井美里氏が note に公開した 2 本の記事[1,2]を対象とする。記事は、口腔用組成物の日本公報 194 件を素材に、①Python による段落単位の JSON 化と事前カウント、②LLM（Claude API）による成分の「役割」の読み分け、③Python による根拠引用の機械照合、④人による抜取確認と定義の修正、という 4 段構成の処理を、費用（約 116.5 ドル）と失敗例を含めて実データで開示した。

評者は、本記事を「生成 AI 時代の特許調査における品質管理の設計図」を示した実務報告として高く評価する。他方、機械照合が保証するのは「根拠が実在すること」までであり、20 件・評価者 1 名の目視確認では分類の品質を統計的に主張できない。また、分類体系（特に「発明の主役」「有効成分」の定義）と独立項の判定には専門的に詰める余地があり、データ取得・API 利用・結果公開に関する規約と権利の確認も記事からは読み取れない。

なお、本記事に対する第三者の公表された論評は、評者の調査範囲（note、X、ブログ等）では見当たらなかった。そこで本評釈では、外部の研究と公的資料に照らして想定される肯定的見解と否定的見解を整理し、そのうえで評者の見解を述べる。

1. 対象記事の書誌と概要

1.1 書誌

項目	内容
著者	酒井美里（スマートワークス株式会社〔Smartworks Inc.〕代表取締役）。セイコーエプソン特許室、エプソンインテリジェンス調査解析部を経て 2005 年に独立[3]。2007 年の第 1 回特許検索競技大会の優勝者[4]
媒体・公開日	note。前編 2026 年 10 月 12 日、後編 2026 年 10 月 13 日（依頼者提供の PDF の表示による。個別記事の URL は未確認）
契機	オンライン講習会の質疑で受けた「公報をどんな方法で AI に読ませているのか。商用 DB の AI 機能や NotebookLM か」という質問への回答の記事化
素材	歯磨剤・洗口液など口腔用組成物の日本公報 194 件。対象成分は糖アルコール（キシリトール、エリスリトール、ソルビトール等）。延べ 991 成分行、表 1,022 表
使用技術	Python（VS Code 上で作成）、Claude API（モデル名・版の記載はない）
公開物	判定結果と根拠引用の一覧を公開サイトで全件公開[5]

1.2 前編の要旨

- **入力と前処理**：入力は商用 DB の CSV（主）、公報 PDF（補足）、RPA による各国特許庁サイトへのアクセス（補足）である。CSV の 1 セルに連結された【0001】【請求項 1】などを段落・請求項ごとに分割し、すべての記載に「住所（ID）」を付けた JSON に変換する。表は CSV に含まれないため PDF から画像で取り出し、1 割強ある横向きの表を回転してから LLM で読み取る[1]。
- **役割分担**：「数える：Python／読む：LLM／確かめる：Python／判断する：人」。同じ入力で結果が揺れては困る処理（件数、文字列の包含、掛け算）は Python に任せる[1]。
- **実験**：糖アルコールの公報内での役割を、A 発明の主役／B 甘味以外の有効成分／C 甘味剤・矯味剤／D 実施例のみ／E 任意成分の列挙／F 配合を制限、の 6 分類（＋判断不可）で読み分けさせ、役割・段落番号・引用原文を出力させる[1]。
- **照合**：Python で「引用文が当該段落に実在するか」「成分名が引用文に含まれるか」「配合量が表の同じセルの値と一致するか」を照合し、NG なら理由を添えて自動で再実行する。194 件は 1 回目 165 件、2 回目 28 件、3 回目 1 件で全件が一致した[1]。
- **教訓**：①事前に Python で数えた延べ 991 成分のうち 76 成分（7.7%）が 1 回目に出力されておらず、先に数えておけば漏れの一部を自動で検出できる。②試走時の照合 NG9 項目は

すべて照合規則側の誤りだった。③原料が水溶液の場合、表の数値はそのまま配合量ではない（例：マルチトール液〔57%〕5.0→約 2.85%）[1]。

1.3　後編の要旨

- **第 1 回改善**：独立項が上位概念「糖アルコール」で、従属項に具体名がある公報まで A（発明の主役）に分類されていたため、役割とは別に「請求項ランク」（1〜5）の軸を設けた。独立項か従属項かは「請求項 N に記載の」という引用から Python で判定した（194 件・2,078 項を 0.06 秒）。A に分類されていた具体名 211 行のうち 65 行がランク 3〜5 に移った。規則だけで直すと 65 行すべてが A から外れたが、LLM で読み直すと B40 行など妥当な振り分けが得られた[2]。
- **第 2 回改善**：著者が 20 件（要注意 6 件＋役割ごとに 2 件ずつ無作為抽出した 14 件）を 20 分で目視し、3 件で LLM と判断が異なった。内訳は、①照合を通過したが役割判断が誤っていた例（マルチトールを A とした例）、②上位概念「糖アルコール」の行を何で判断するかが未定義だった例、③A〜F のいずれにも当たらない「課題の原因となる成分をあえて使う」型、である。定義を直して上位概念が記載された 317 行を読み直すと、143 行の役割が変わり、判断不可は 41 行から 16 行に減った[2]。
- **時間と費用**：API 費用は合計約 116.5 ドル、実行時間は約 1 時間 36 分。人の確認は、AI の判定と根拠段落がある状態で 1 公報約 1 分、ゼロから読むと 2〜3 分だった。設計と定義修正の時間は表に出ないが作業価値の大部分はそこにあり、LLM の費用は「渡す本文の量×呼び出しの数」で決まるので読ませる項目は最初に決めておく方が安い、と述べる[2]。

1.4　記事中の数値の検算

記事に示された数値を評者が再計算した結果を表 1 に示す。主要な数値は整合しており、記事の報告は内部的に一貫している。

表 1　記事中の数値の検算

項目	記事の記載	評者の検算	判定
照合一致の内訳	1 回目 165＋2 回目 28＋3 回目 1	計 194 件	整合
出力漏れ率	76／991＝7.7%	7.67%	整合
濃度換算	5.0×0.57≒2.85%	2.85%	整合
第 1 回改善の再分類	A3・B40・C6・D6・E8・判断不可 2	計 65 行	整合

項目	記事の記載	評者の検算	判定
目視の標本	要注意 6＋無作為 14＝20 件	7 区分（A〜F＋判断不可）×2 件＝14 件	整合
費用合計	約 116.5 ドル	116.54 ドル	整合
第 2 回改善の見込額	行数が 3 分の 1 なので約 8 ドル	役割読み分け 31.85 ドルの 3 分の 1 は約 10.6 ドル（行数比 317／991 なら約 10.2 ドル）	根拠を再現できない（軽微）
目視での不一致率	3／20	15%（95%信頼区間：Wilson 法で約 5.2〜36.0%）	区間が広い（4.3 参照）

2. 総合評価

論点ごとの肯定的見解・否定的見解と評者の見解を表 2 にまとめる。詳細は第 3 章（肯定的見解）、第 4 章（否定的見解）、第 5 章（評者の見解）で述べる。

表 2　論点別の評価の一覧

論点	肯定的見解	否定的見解	評者の見解
分業設計	揺れては困る処理をコードに寄せる設計は外部研究の知見と整合する	照合規則も人が作るため誤り得る。分業の境界は固定できない	支持。規則側の誤りを前提に回帰テストを組み込んだ点まで評価する
根拠照合	段落 ID で機械照合でき、監査可能性が高い	確認できるのは実在まで。引用が判断を支持するかは保証しない	支持。ただし実在と妥当性の間に「支持性」の関門が必要
事前カウント	出力漏れの一部を自動で検出できる	辞書にない成分の漏れは検出できない	実務上重要な工夫
品質評価	要注意公報を含めた厳しめの評価	20 件・評価者 1 名・非無作為。誤り率の上限を示せない	品質の主張には不十分
分類体系	請求項ランクの分離は適切な修正	A・B の定義に特許法・薬事上のずれ。類型の欠落	定義の再設計を推奨
再現性・費用	費用・失敗まで開示し透明	モデル名・プロンプト・母集団の条件が不明	限界費用として読むべき
規約・権利	（記事では言及なし）	J-PlatPat、商用 DB、API、公開の各確認が不明	業務利用の前に要確認

3. 肯定的見解

前述のとおり、本記事への第三者の公表論評は見当たらなかった。以下は、外部の研究・公的資料に照らして、本記事の手法を肯定的に評価できる論拠である。

3.1 「LLM に全部やらせない」分業は外部研究の知見と整合する

法務分野では、検索拡張生成（RAG）を用いる商用リーガルリサーチ AI でも 17〜33%の割合でハルシネーションが生じたとする事前登録型の実証研究がある[6]。同研究は、ハルシネーションに「出典が記述を支持していると誤って主張する」場合を含めて定義している[6]。汎用 LLM に無作為に選んだ米国連邦裁判所の判例について具体的で検証可能な質問をすると、58〜88%の割合でハルシネーションが生じたとの報告もある[7]。医療分野でも、LLM の回答の 50〜90%が、引用した出典によって完全には支持されていなかった[8]。さらに、temperature を 0 に設定しても、同一入力に対する出力が安定しないことが実測されている[9]。

これらの知見は、「根拠付きで答えさせれば正しくなる」とは言えないこと、そして出力の揺れを前提に設計すべきことを示している。本記事が、揺れては困る処理（計数・包含判定・計算）を Python に寄せ、LLM には「読む」作業だけを任せ、最終判断を人に残す構成をとったことは、こうした知見と整合する合理的な設計である。評者は以前、生成 AI の知財業務での活用方法を、①そのまま利用する、②社内外のデータと連携させる、③外部ベンダーのサービスを利用する、の 3 類型に整理し、特許調査では分類作成支援や特許読み込み支援が有望であると述べた[10]。本記事は②の類型を、検証工程まで含めて具体化した実例と位置づけられる。

3.2 段落 ID による「住所」付与は特許文献の引用作法と一致する

公報を段落・請求項の単位に分割し、すべての記載に ID を振る設計によって、LLM の主張は「段落 0025 にこう書いてある」という検証可能な形になる。特許実務では、拒絶理由通知でも鑑定でも段落番号や請求項番号で引用箇所を特定するのが通例である。この単位で照合ログが残れば、調査結果を第三者が追試・監査できる。記事が「ID がなければ、確認できるのは CSV の『実施例』セルに含まれるという程度の粒度になる」と指摘する点[1]は、AI 出力の監査可能性を左右する本質を突いている。

3.3 事前カウントは「漏れ」の検証という難題への部分的な解答である

特許調査では、見つけたものが正しいか（適合率）よりも、見落としがないか（再現率）を検証する方が難しい。本記事は、LLM に読ませる前に対象成分の出現数を Python で数えておき、出力と突き合わせることで、1 回目の出力漏れ（76 成分、7.7%）を機械的に検出し、再実行で

回収した[1]。著者は過去に「正しさは検証できても、漏れは検証できない」と述べていたが[1]、少なくとも辞書で定義できる範囲では漏れも検証できることを実データで示した。この意義は大きい。

3.4 「照合する側を疑う」姿勢と回帰テストは運用の成熟度を示す

試走時の照合 NG9 項目がすべて照合規則側の誤りであったという報告[1]は、AI 出力の検証で見落とされやすい「検証器自身の誤り」を正面から扱っている。規則を直すたびに記録済みのデータで照合をやり直し、規則に起因する NG がゼロになることを確かめてから本番に流す運用は、ソフトウェア開発における回帰テストに相当する。API 費用をかけずに検証の精度を高める工夫として、実務的な価値が高い。

3.5 表の処理と濃度換算は化学系調査の実務課題に踏み込んでいる

化学系の公報では配合量の多くが表に記載されており、しかも原料が水溶液や市販品の場合、表の数値はそのまま有効量を示さない。本記事は、1,022 表を画像から読み取り、注記から濃度を読み取る作業（LLM、根拠引用付き）と掛け算（Python）とを分けた[1,2]。配合量に基づく権利範囲の判断や侵害予防調査では、この種の換算漏れが致命的な誤りになり得るため、実務上の価値は大きい。

3.6 費用・失敗・結果を開示した透明性

本記事は、API 費用を工程別に開示し（合計約 116.5 ドル）、「設計時間と定義を直す作業時間にこそ作業価値の大部分がある」と明言し、見込み違いの費用（約 8 ドルの見込みに対して実際は約 20 ドル）まで記している[2]。判定結果と根拠も公開サイトで全件公開している[5]。生成 AI の業務活用に関する記事が成功例の紹介に偏りがちななかで、失敗と修正の過程を定量的に示した点は、実務家にとって再利用価値の高い知見である。

3.7 「どのツールか」から「どう設計するか」への問いの組み替え

質問者は商用 DB の AI 機能や NotebookLM といった「ツール」を尋ねたが、記事は「大事なのはツール選びよりも設計」と答えた[2]。特許庁の調査では、調査対象の商用 DB の多くが既に大規模言語モデルを類似検索・概念検索のエンジン等に用いていることが確認されている[11]。ツールが横並びになりつつある状況で、何を数え、何を照合し、どこで人が判断するかという設計が品質を決めるという指摘は、ツール選定に関心が集まりがちな現場への有益な問題提起である。

4. 否定的（批判的）見解

4.1 照合の通過は分類の妥当性を保証しない

記事自身が認めるとおり、Python の照合で確認できるのは「その段落が実在し、成分名を含む」ことまでである[2]。特許 5797243 の例では、マルチトールを A（発明の主役）とした誤判断が、引用の実在性の照合を通過した[2]。引用の評価に関する研究では、引用文が実在することと、引用文が主張を「支持」しているか（含意）とを区別し、後者の自動判定には教師あり言語モデルや LLM が用いられている[8]。本記事の照合は前者にとどまり、後者は人の抜取確認に全面的に依存している。

著者は確認を「第一関門＝指摘した記載が実在するか」「第二関門＝その指摘が妥当か」に分けている[2]。評者は、その間に「引用文が付与された役割を支持しているか」という中間の関門を置き、そこを別モデル等の機械で一次判定し、判定が割れた行を人に回す設計が有効だと考える（第 5.2 節）。ただし、LLM を判定者とする手法にも誤りや偏りがあり、万能ではない点には留意が必要である。

4.2 「照合に通るまで再実行」は、照合に通る誤りを選び出すおそれがある

NG のたびに理由を添えて再実行させる方式は、照合ルールという指標に適合する出力を選抜する。そのため、ルールで捕捉できない種類の誤り（役割の取り違え）を含みながら照合を通る出力が残りやすいという構造上のリスクがある。指標を目標にすると指標としての意味が失われるという、いわゆるグッドハートの法則に類する問題である。また、LLM の出力は同じ条件でも揺れるため[9]、1 回目で一致した 165 件も、再実行すれば役割判断が変わる可能性がある。記事は再実行時の判断の安定性を測っていない。同じ行を複数回読ませて多数決をとる自己整合性（self-consistency）の手法[12]を応用し、票が割れた行を「要確認」とすれば、追加の人手なしに不安定な判断を抽出できる。

4.3 20 件・評価者 1 名の目視確認では品質を統計的に主張できない

目視で判断の相違が見つかったのは 20 件中 3 件（15%）であるが、二項分布に基づく 95%信頼区間は、Wilson 法で約 5.2〜36.0%、Clopper-Pearson 法で約 3.2〜37.9%と広い（評者の計算）。加えて、次の制約がある。

- 標本は要注意の 6 件を意図的に含んでおり、母集団の無作為標本ではない。
- 分類の単位は「行」（991 行）であるのに、抽出の単位は「件」（公報）である。1 件に複数の行がある場合の扱いが明らかでない。

・評価者は著者 1 名であり、評価者間の一致度（κ 係数等）が測られていない。

・評価者が分類定義の設計者でもあるため、設計者の想定に沿った判断になりやすい。

記事は「全数チェックよりは厳しめの数字」と注記しているが[2]、仮に偏りがその方向にあるとしても、誤り率の上限は示せない。顧客に「品質の報告」を提供するのであれば[2]、層化無作為抽出による 50〜100 行規模の標本、2 名以上による独立判定と一致度の算出、不一致の合議による正解データの作成、定義の修正に用いていない別の公報群での検証、という手順が望ましい。

4.4　分類体系には特許法・薬事の観点から詰めるべき点がある

（1）「発明の主役」の定義が二つの基準を混在させている

分類 A は「請求項で必須とされる、または課題解決の中心とされる」と定義されている[1]。しかし、請求項には出願人が「発明を特定するために必要と認める事項のすべて」を記載するのであって[13]、請求項に記載された成分が課題解決の中心であるとは限らない。性質の異なる二つの基準を OR で結んだことが、後編で「主役」が多すぎる結果を招いた一因と考えられる。請求項ランクを別の軸として分離した後編の修正[2]は適切であるが、分類 A 自体も「請求項上の位置づけ」と「課題解決上の位置づけ」の 2 軸に分ける方が、判断の再現性は高まる。

（2）「有効成分」という用語が薬事上の意味とずれている

分類 B は「甘味以外の有効成分（う蝕予防、保湿、安定化など）」とされ、記事は湿潤剤として配合されるキシリトールを B の例に挙げている[1]。しかし、医薬部外品である薬用歯みがき類では、「有効成分」は承認基準の別表に掲げられた成分を指す用語である。令和 3 年の薬用歯みがき類製造販売承認基準の別表 1〜3 には、キシリトール、ソルビトール、エリスリトール、マルチトールなどの糖アルコールは掲げられておらず、湿潤剤などは添加剤として別に扱われる[14]。口腔用組成物の公報を読む実務家、特に薬事に通じた読者に誤解を与えるため、「甘味以外の機能目的成分」などに改称すべきである。

（3）「課題の原因成分を前提とする改良発明」の類型が欠けている

特許 2021069317 のように、課題の原因となる成分の使用を前提とし、その欠点を打ち消す発明は、A〜F のいずれにも当てはまらなかった[2]。組成物の発明では、ある成分に起因する不具合（変色、沈殿、苦味、効果の減殺など）を別の手段で解消する改良発明は典型的な類型であり、分類の設計段階で想定しておくべきであった。

（4）結果を見てから定義を変え、同じデータで再評価している

上位概念の判断基準は目視の結果を見た後に修正され、同じ 194 件で再評価された[2]。定義の修正そのものは健全な改善であるが、修正後の精度を同じデータで測れば推定は楽観的になる。また、317 行中 143 行（45%）の役割が定義変更で動いたことは、分類結果が定義に強く依存すること、つまり定義を固めない限り結果が安定しないことを示している。

4.5 「請求項 N に記載の」による独立項・従属項の判定には限界がある

特許庁の審査ハンドブックは、他の請求項を引用して記載する「引用形式請求項」と「独立形式請求項」とを、記載表現が異なるだけで同等に扱われるものとしている。そのうえで、引用形式は、引用される請求項に係る発明と同一か否かに関わらず用いることができるとし、請求項 1 のボールベアリングを引用して「その製造方法」を記載する、カテゴリーの異なる請求項を例示している[15]。したがって、「請求項 1 に記載の組成物を含有する歯磨剤」や「請求項 1〜3 のいずれか 1 項に記載の組成物の製造方法」のような実質的に別発明の請求項も、引用表現を検出するだけでは「従属項」と判定されてしまう。

加えて、複数の請求項を択一的に引用する表記（「請求項 1 又は 2」「請求項 1 から 3 のいずれか一項」など）には揺れがあり[16]、規則の網羅性が判定精度を左右する。独立項にマーカッシュ形式で具体名が含まれる場合（選択肢の一つとして記載）をランク 1 とするのか、従属項の選択肢（ランク 3）と区別しないのかも、記事からは明らかでない。0.06 秒という処理速度は魅力的だが、カテゴリーの異なる引用形式請求項を検出する規則を加えることが望ましい。

4.6 母集団と再現性に関する情報が不足している

194 件の抽出条件（検索式、対象期間、公報の種別）は記事に示されていない。記事中では、「特許 4959006」のような登録番号の形式と、「特許 2014214110」「特許 2026065720」のような公開番号の形式の番号が、いずれも「特許」と表記されている[1,2]。公開公報と特許公報が混在している可能性がある（評者の推認）。同一出願の公開公報と特許公報が双方含まれていれば、件数や役割の分布に重複の影響が出る。さらに、使用したモデル名・版、プロンプト、temperature 等の設定も示されていないため、第三者による追試や、費用の他案件への当てはめはできない。

4.7　データ取得・API 利用・結果公開の規約と権利の確認が明らかでない

（1）J-PlatPat

INPIT は、データの単純な収集を目的とした大量のダウンロードや、プログラムによる定期的な自動データ収集（ロボットアクセス）を行わないよう求めている[17]。記事が補足データとして挙げる「RPA で各国特許庁にアクセス」[1]が、日本について J-PlatPat を対象とするのであれば、この利用条件との関係を説明する必要がある。特許庁は、一括ダウンロードサービス[18]や特許情報取得 API[19]という正規の大量取得手段を用意しており、自動化を前提とするならこれらを使うのが筋である。各国特許庁サイトの利用条件は個別に確認が必要である。

（2）商用 DB

商用 DB から出力した CSV を外部の生成 AI の API に投入することが利用契約上許されるかは、DB や契約ごとに異なり得る。評者の調査では一般化できる情報は得られなかったが、業務で用いる前に書面で確認すべき事項である。

（3）API への送信と守秘

公報そのものは公開情報であるが、顧客の調査テーマ、母集団、判定基準（プロンプト）は秘密情報に当たり得る。Anthropic の Privacy Center は、API 利用者の入出力を、原則として受領・生成から 30 日以内にバックエンドから削除する（ゼロデータ保持契約等の例外がある）と説明している[20]。また、同社の商用利用規約は、顧客コンテンツをモデルの学習に用いない旨を定めているとされる[21]。委託契約の守秘条項とこれらの条件を照らし合わせ、必要に応じてゼロデータ保持の適用可否を確認することが望ましい。

（4）公開サイトへの掲載と著作権

INPIT は、公報に掲載された明細書等の文章や図面には通常出願人等の著作権があり、転載には許諾が必要となる場合がある旨を注意喚起している[17]。文化審議会著作権分科会法制度小委員会の「AI と著作権に関する考え方について」は、情報解析のための利用（著作権法 30 条の 4）と、既存の著作物の創作的表現を出力に含める利用とを区別して論じている[22]。根拠として短い原文を示す本件の公開は、引用（同法 32 条）や、情報解析の結果の提供に付随する軽微利用（同法 47 条の 5）として許容される余地が大きいと考えられる。ただし司法判断の蓄積はなく、評価は確定していない。

4.8　費用・時間の比較は「限界費用」として読むべきである

約116.5ドル、1公報約1分という数字は、設計済みの仕組みがある状態での限界費用であり、記事自身が述べるとおり設計と定義修正の時間は含まれていない[2]。また、Claude API の料金は入力トークンと出力トークンに課金され、バッチ処理の割引やプロンプトキャッシュも用意されている[23]。「渡す本文の量×呼び出しの数」という記事の整理[2]は、全文を読ませて短いラベルを返させる処理の近似としては妥当である。ただし、第 2 回改善のように公報単位で全文を再投入する場合は、キャッシュやバッチ処理で費用を圧縮できた可能性がある。表 1 のとおり、約 8 ドルという見込額の算定根拠も記事からは再現できない。

5. 評者の見解

5.1　本記事の位置づけ

本記事の最大の貢献は、生成 AI を特許調査に用いる際の品質管理を、心構えではなく工程として設計し、その工程がどこで機能し、どこで機能しなかったかを実データで示した点にある。「数える：Python、読む：LLM、判断：人」という分業は単純であるが、単純であるがゆえに組織内で共有・監査しやすい。特許調査の経験が豊富な著者が、プログラミングの初学者として Claude と対話しながら構築したという事実[2]も、知財部門が外部ベンダーに全面依存せず、自前の検証の仕組みを持ち得ることを示しており示唆的である。

他方で、本記事が到達したのは「根拠の実在が機械的に保証された分類結果」であって、「品質が統計的に保証された分類結果」ではない。この差は、社内の探索的な分析では許容できても、侵害予防調査や鑑定、顧客への品質報告では決定的になり得る。

5.2　「二つの関門」から「三つの関門」へ

評者は、著者の「二つの関門」を次の三段階に拡張して運用することを提案する。

表 3　評者が提案する三段階の確認

関門	確認する内容	担い手と方法
第一関門（実在性）	引用文が当該段落に存在するか。成分名・数値が一致するか	Python で全件を機械判定（現行どおり）
第二関門（支持性）	引用文が付与された役割を支持しているか	別モデル等で全件を一次判定。複数回の読み直しで判断が割れた行とあわせて「要確認」に振り分け
第三関門（妥当性）	役割の判断が発明全体の理解として妥当か	層化抽出した行と要確認行について、複数の人が独立に判断し一致度を記録

このように機械の役割を「実在の確認」から「人が見るべき行の選別」へ広げれば、人の確認時間を増やさずに品質の主張を一段強くできる。

5.3　実務への示唆

表 4　課題と推奨策

課題	推奨策	期待される効果
照合が表層的	支持性の判定（別モデル等）と、自己整合性による要確認行の抽出を加える	照合を通る誤りの低減、人の確認対象の絞り込み
品質評価の根拠	層化無作為抽出で 50〜100 行、2 名の独立判定、κ 係数と信頼区間の併記、検証用データの分離	顧客に示せる品質指標
分類の定義	A を 2 軸に分解、B を「機能目的成分」に改称、「課題原因成分の克服」型を追加、定義は試走で確定させる	判断の再現性向上、再処理費用の削減
独立項の判定	カテゴリーの異なる引用形式請求項の検出、択一的引用の表記の網羅	請求項ランクの精度向上
再現性	モデル名・版、プロンプト、設定、実行日、母集団の抽出条件を記録・開示する	追試と費用見積もりが可能になる
規約・権利	J-PlatPat は一括ダウンロードや API を使う。商用 DB 契約と API の守秘条件を書面で確認する。公開時は引用の要件を満たす	コンプライアンス上のリスク低減

5.4　今後の展望

特許庁は、令和 11 年（2029 年）1 月に予定する J-PlatPat の刷新において、AI の活用やパテントマップ作成機能の実装などを検討している[24]。公的サービスにも商用 DB にも AI 機能が標準で備わっていくなかで、差がつくのはツールそのものではなく、出力をどう数え、どう照合し、どこで人が判断するかという検証の設計である。本記事は、その設計を個人の実務家が公開の場で示した先行例として、知財部門と調査会社の双方にとって参照価値が高い。上記の補強がなされれば、生成 AI 時代の「調査品質」の標準的な枠組みに発展し得ると評者は考える。

6. 結語

本記事は、「公報をどう AI に読ませているか」という素朴な質問に、ツール名ではなく工程と数字で答えた。その誠実さこそが本記事の価値である。一方で、機械照合の射程、品質評価の

統計的な限界、分類定義の専門的な詰め、規約と権利の確認という 4 点は、業務として提供する段階では避けて通れない。著者は記事の末尾で、「何を数え、何を照合し、どこで人が判断するかを決める部分」こそが特許調査の品質管理そのものだと述べている[2]。そうであれば、次の課題はその「決め方」自体を検証可能にすることである。

【注記】本評釈は、依頼者から提供された記事の PDF 版・Word 版に基づく。評者は公開サイトの内容を直接確認していない。確認済みの事実と評者の推論・意見は本文中で区別した。法的評価は確定したものではない。URL の閲覧日は 2026 年 11 月 8 日である。

参考文献

[1] 酒井美里「講習会 Q&A（前編）『どうやって公報を AI に読ませてますか？』に実データで答えます」note（Smartworks Inc. / 特許調査）、2026 年 10 月 12 日公開。依頼者提供の PDF・Word 版に依拠。著者ページ：https://note.com/sakaimisato （個別記事 URL は未確認）

[2] 酒井美里「講習会 Q&A（後編）『どうやって公報を AI に読ませてますか？』に実データで答えます」note（Smartworks Inc. / 特許調査）、2026 年 10 月 13 日公開。依頼者提供の PDF・Word 版に依拠。著者ページ：https://note.com/sakaimisato （個別記事 URL は未確認）

[3] 一般財団法人日本特許情報機構（Japio）掲載「スマートワークス株式会社 代表取締役 酒井美里」プロフィール、2017 年。https://japio.or.jp/hyousyou/files/2017/17-a-01.pdf

[4] 情報科学技術協会（INFOSTA）「INFOSTA 会員限定 新春セミナー 元気なインフォプロを目指して」案内、2009 年 1 月 16 日。https://www.infosta.or.jp/seminar/20090116.pdf

[5] 酒井美里（Smartworks Inc.）「甘味成分の役割読み分け」公開サイト（記事中のリンク表示による）。https://deluxe-pothos-5b9c2f.netlify.app （評者は内容を直接確認していない）

[6] V. Magesh, F. Surani, M. Dahl, M. Suzgun, C. D. Manning, D. E. Ho, "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools," Journal of Empirical Legal Studies, Vol.22, No.2, pp.216–242 (2025). doi:10.1111/jels.12413（arXiv:2405.20362）。https://arxiv.org/abs/2405.20362

[7] M. Dahl, V. Magesh, M. Suzgun, D. E. Ho, "Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models," Journal of Legal Analysis, Vol.16, No.1, pp.64–93 (2024). doi:10.1093/jla/laae003。https://arxiv.org/abs/2401.01301

[8] K. Wu, E. Wu, K. Wei, A. Zhang, A. Casasola, T. Nguyen, S. Riantawan, P. Shi, D. E. Ho, J. Zou, "An automated framework for assessing how well LLMs cite relevant medical references," Nature Communications, Vol.16, Article 3615 (2025). doi:10.1038/s41467-025-58551-6。https://pmc.ncbi.nlm.nih.gov/articles/PMC12003634/

[9] B. Atil et al., "Non-Determinism of 'Deterministic' LLM Settings," arXiv:2408.04667 (v5). https://arxiv.org/abs/2408.04667

[10] 萬秀憲「生成 AI の知財業務での活用」知財管理 Vol.74 No.7、pp.828–842（2024 年）（拙稿）。紹介頁：https://yorozuipsc.com/blog/7ai5350777

[11] 特許庁「令和 5 年度 特許情報に係る商用データベースの機能水準に関する調査報告書」（令和 5 年度特許庁請負事業、NTT アドバンステクノロジ株式会社）、令和 6 年 3 月。
https://www.jpo.go.jp/resources/report/sonota/service/document/index/2023hokokusho.pdf

[12] X. Wang, J. Wei, D. Schuurmans, Q. Le, E. Chi, S. Narang, A. Chowdhery, D. Zhou, "Self-Consistency Improves Chain of Thought Reasoning in Language Models," arXiv:2203.11171 (ICLR 2023).
https://arxiv.org/abs/2203.11171

[13] 特許法（昭和 34 年法律第 121 号）第 36 条第 5 項。e-Gov 法令検索：https://laws.e-gov.go.jp/law/334AC0000000121

[14] 厚生労働省医薬・生活衛生局長通知「薬用歯みがき類製造販売承認基準について」（令和 3 年 6 月 28 日薬生発 0628 第 13 号）別紙・別表 1〜3。https://www.mhlw.go.jp/content/000797783.pdf

[15] 特許庁「特許・実用新案審査ハンドブック」第 II 部第 2 章 2202「請求項の記載形式―独立形式と引用形式―」（2015.10）。
https://www.jpo.go.jp/system/laws/rule/guideline/patent/handbook_shinsa/document/index/02.pdf

[16] 特許庁「特許・実用新案審査基準」第 II 部第 2 章第 5 節「特許請求の範囲の記載に関する委任省令要件（特許法第 36 条第 6 項第 4 号）」。
https://www.jpo.go.jp/system/laws/rule/guideline/patent/tukujitu_kijun/document/index/02_0205.pdf

[17] 独立行政法人工業所有権情報・研修館（INPIT）「J-PlatPat 利用上のご案内」。
https://www.inpit.go.jp/j-platpat_info/guide/j-platpat_notice.html （調査段階の取得内容に依拠。評者は本稿執筆時に原文を再取得できておらず、要確認）

[18] 特許庁「特許情報の一括ダウンロードサービスについて」（令和 7 年 7 月 24 日）。
https://www.jpo.go.jp/system/laws/sesaku/data/download.html

[19] 特許庁「特許情報取得 API 利用ハンドブック（第 2.0 版）」。
https://www.jpo.go.jp/system/laws/sesaku/data/document/api-provision/api_handbook_v2.0.pdf

[20] Anthropic Privacy Center, "How long do you store my organization's data?"（2026 年 11 月 8 日閲覧）。
https://privacy.claude.com/en/articles/7996866-how-long-do-you-store-my-organization-s-data

[21] Anthropic, "Commercial Terms of Service"（2025 年 6 月 17 日発効とされる）。
https://www.anthropic.com/legal/commercial-terms （評者は原文を直接確認しておらず、二次資料における引用に依拠。要確認）

[22] 文化審議会著作権分科会法制度小委員会「AI と著作権に関する考え方について」（令和 6 年 3 月 15 日）。
https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_07/pdf/94024201_01.pdf

[23] Anthropic, "Pricing"（Claude Docs、2026 年 11 月 8 日閲覧）。https://docs.claude.com/en/docs/about-claude/pricing

[24] 特許庁「令和 7 年度 特許情報提供サービスの現状と今後に関する調査事業 報告書」（令和 7 年度特許庁請負事業、NTT アドバンステクノロジ株式会社）、令和 8 年 3 月。
https://www.jpo.go.jp/resources/report/sonota/service/document/index/2025hokokusho.pdf