Gemini 4 Carbon「AA Intelligence Index v5 首位」の検証と知財業務への影響予測

2026 年 10 月 29 日時点

Claude Opus 5.5

凡例

【事実】＝一次ソースまたは報道で確認できた事項。【分析】【予測】＝本稿による分析・推測。
本文中の上付き番号は文末の参考文献に対応する。参考文献の記号：無印＝本稿作成時に本文を取得して確認／※＝調査段階または検索結果の抜粋で確認（本文全体は再確認していない）／【未検証】＝本文未取得・内容未確認。

要旨

- **前提は真【事実】**：Artificial Analysis（以下 AA）の記事[1]と Google 公式ブログ[2]（いずれも 2026 年 10 月 28 日）で、Gemini 4 Carbon が Intelligence Index v5 で 59 点・首位であることを確認した。2 位は Claude Opus 5.5（58）、3 位は GPT-6 Astra（57）[3]。
- **限定提供【事実】**：現時点の利用者は Fairwind Program 経由の一部サイバー防御者のみで、開発者・企業・一般向けの提供日は示されていない[2]。AA も「公開 API 無し」と表示している[3]。
- **首位の中身【事実】**：9 評価中 4 つで 1 位だが、コーディングとナレッジワークでは首位ではない[1]。首位との差は 1 点で、不確実性の推定も示されていない[3]。物差し（v5）自体も公開 2 日目である[4]。
- **知財への含意【予測】**：長文脈・長時間エージェント・リーガルリサーチ系の強さは、先行技術調査・無効資料調査・IP ランドスケープ（IPL）の自動化に追い風となる。一方、AA-Omniscience のハルシネーション率は 4%から 11%に上昇した[1]。In re Mitchell 事件[5]が示すとおり、引用の原典照合は人が担うという前提は崩せない。今は乗り換えず、自社評価セットとガバナンスの整備を優先すべきである。

1. 記事内容の確認（AA 記事）【事実】

AA は 2026 年 10 月 28 日、Gemini 4 Carbon（high reasoning）が Intelligence Index v5 で 59 点の首位になったと公表した[1]。主な内容は表 1 のとおり。

表 1　AA 記事の主な内容

項目	内容
総合スコア	59 点（Gemini 4 Argon（high）の 50 点から+9）。v5 公開時（10 月 27 日）の首位は Claude Opus 5.5（max）58、次いで GPT-6 Astra（max）57、GPT-6.1 Sol（max）55[4]
首位の評価	9 評価中 4 つ：HLE-Diamond、Terminal-Bench-Science、AA-Omniscience、AutomationBench-AA[1]
コーディングエージェント	Coding Agent Index 59 点（Antigravity CLI 経由、Argon は 53）。首位は Claude Code＋Opus 5.5（max）の 62 点。59 点は Claude Code＋Sonnet 5.5、Muse Code＋Muse Wonder Full と同点[1,3]
同指数の内訳	Terminal-Bench 4.0 64%、FrontierCode v1.1 51%（いずれも 2 位。FrontierCode の首位は Opus 5.5 の 56%）、SWE-Atlas-QnA-AA 64%（首位は Muse Code の 69%）[1,3]
ナレッジワーク	AA-Briefcase 1609 Elo（Argon 比+126）。Grok 4.7（xhigh）1638、Muse Wonder（max）1659、Kimi K3.1 Pro 1678 を下回る。改善幅は Presentation +293（1598）、Analytical Quality +120（1671）[1]
コスト	標準価格換算で 1 タスク$7.76。Opus 5.5（max）の$10.04 より約 23%安く、GPT-6 Astra（max）の$6.25 の約 1.2 倍。出力は 1 タスク 12.3 万トークンで、Opus 5.5（17.3 万）より 29%少ない[1]
知識・ハルシネーション	AA-Omniscience v1.1 で 54 点（Opus 5.5 は 48）。正答率は 58%（Argon 比+10pt）だが、ハルシネーション率は 4%→11%に上昇。それでも Argon に次ぐ低さで、Opus 5.5 と Astra はいずれも 50%[1]
仕様・価格	コンテキスト長 1M トークン。入力はテキスト・画像・動画・音声、出力はテキスト。標準価格は 100 万トークン当たり入力$4／出力$20、現在 50%のローンチ割引、キャッシュ入力は 95%引き[1]
速度	出力約 92〜102 トークン/秒、初回応答まで約 97.5〜100 秒（比較対象の中央値 3.6 秒）とされる[6,7]

ローンチ割引の終了日について、AA は「確定した終了日はない」とする[1]一方、The Decoder は「2026 年末まで」と報じており[8]、情報が食い違っている。

2. 一次ソース（Google 公式発表）の確認【事実】

- **発表：**2026 年 10 月 28 日、Koray Kavukcuoglu（Google DeepMind SVP 兼 Chief AI Architect）名義のブログ「Gemini 4 Carbon: our next frontier model」で発表。冒頭で AA Intelligence Index v5.0 で 59 点・首位であることを明記している[2]。
- **提供形態：**今週から一部のサイバー防御者に提供し、Fairwind Program を通じて信頼できるパートナーへ段階的に広げる。ガードレールを改善したうえで、開発者・企業・一般消費者向けに「できるだけ早く」提供するとし、日付は示していない。米国政府の自主的なリリース前モデルテストにも参加している[2]。
- **価格：**導入価格は Argon と同じ入力$2／出力$10（100 万トークン当たり）、キャッシュ入力は 95%引き。導入期間後は$4／$20（脚注）[2]。Argon の導入価格は 11 月 30 日終了と開発者向け文書に明記されているのに対し、Carbon の終了日は発表されていない[3]。
- **サイバー系ベンチ（ガードレール無しで評価）：**SEC-bench Pro 88.2%（脆弱性発見）、ExploitGym 54.8%（悪用検証）、CWE-bench 71.0%（修正）[2]。
- **その他の自社公表値：**WebDev Arena で+89 Elo の 4 位（1768）。Terminal-Bench 4.0 での不正成功率 0.0%（不正試行 0.4%）。Gray Swan の間接プロンプトインジェクション攻撃成功率は 0.7%→0.0%。社内試験で根拠のない主張を約 2 分の 1 に、重大なミスアライメント警告を約 80%削減し、過剰実行を 26%、支援不足を 29%減らしたとする。セキュアコード生成の成功率は Argon の 2 倍超、Chrome チームの評価では他社フロンティアモデルの 10 倍超[2]。比較相手の社名は非開示で、第三者は検証できない[3]。
- **Google のベンチ表：**17 行中 15 行で首位。例外は FrontierSWE 2（Carbon 61.2%、Astra 65.5%、Opus 62.3%）と OSWorld-2.1 オフライン部分集合（Carbon 81.5%、Opus 87.2%）。長文脈の GraphWalks（256K〜1M トークン）は 96.2%で、Astra の 71.8%、Opus の 66.8%を大きく上回る[3]。ただし GraphWalks は全モデルとも Google 自身が計測したとされる[9,10]。
- **未公開事項：**技術レポート、パラメータ数、アーキテクチャは公表されていない[3]。10 月 28 日時点で Google のモデルカード一覧に Carbon はなく[11]、Gemini API のリリースノートにもモデル ID はない[9]。評価手法を説明する PDF があるとの報道はある[9]が、本稿では確認していない。

- **「Carbon」の位置づけ【未確定】：**Google は「次のフロンティアモデル」「最強のモデル」と呼び、Argon を土台にしたと説明しているが、Pro／Flash といった階層との関係は示していない[2]。Business Insider は 10 月 9〜10 日、社内文書に Argon・Barium・Carbon の名があり、公開版 Argon は Barium-B だったと報じた（原記事は未確認）[12]。「Carbon＝Ultra」説は X 上の未確認情報である[13]。【分析】価格が Argon と同一であることから、上位の階層というより「Argon に続く次のチェックポイント」と読むのが妥当と考える（推測）。

3. AA Intelligence Index v5 の構成と限界

3.1 構成【事実】

v5 は 2026 年 10 月 27 日に公開された。2026 年 1 月の v4 以来最大の改定で、指標の 60%が非公開テストまたは非公開解答に基づく[4]。

表 2　Intelligence Index v5 の構成

分野（比重）	評価（比重）
Knowledge Work（25%）	AA-Briefcase 15%、AutomationBench-AA 5%、GDP.pdf 5%
Coding（25%）	FrontierCode 2.0 15%、Terminal-Bench 4.0 10%
Science（25%）	Terminal-Bench-Science 15%、CritPt v2 hard subset 10%
General（25%）	AA-Omniscience v1.1 15%（正答 10%＋非ハルシネーション 5%）、HLE-Diamond 10%

- **変更点：**FrontierCode 2.0 と Terminal-Bench-Science を追加し、HLE を HLE-Diamond に置き換え、AA-Omniscience と CritPt を改訂した。GDPval-AA、SciCode、AA-LCR（長文脈推論）は外した[4]。
- **性質：**推論努力（effort）設定の影響が大きくなった（Opus 5.5 は low で 35 点、max で 58 点、コストは 10 倍）。1 タスク当たり出力トークンの中央値は 3.5 万から 5.9 万に増え、オープンウェイトとの差は 3 点から 12 点に広がった（最上位は Kimi K3.1 Pro の 46 点、重みは未公開）[4]。

3.2 限界・批判

- 【事実】AA-Omniscience の「ハルシネーション率」は「不正解÷（不正解＋部分正解＋無回答）」であり、「全回答の 11%が捏造」という意味ではない[3]。

- 【事実】Nathan Lambert は、AA がラボから収益を得る事業構造であることや、テストの非公開化でコミュニティがベンチの不具合を見つけにくくなる点を指摘する。同時に、最前線の評価の正確さでは AA や Arena を信頼するとも述べている[14]。
- 【事実】Hacker News では「1 点差に意味はあるのか」「米国ラボ寄りではないか」との疑問が出た[15]。Carbon のスレッドでは、v5 公開翌日に Google が首位となったタイミングへの皮肉も見られた[16]。
- 【分析】知財の観点では、v5 に長文脈専用の評価（AA-LCR は除外）も、多言語や図面理解の評価も含まれない点が最大の限界である。「日本語の特許公報を大量に読み、正確に対比できるか」は、この指標からは分からない。

4. 競合モデルとの比較

表 3　AA Intelligence Index v5 における主要モデル

モデル	AA v5	主な特性
Gemini 4 Carbon（high）	59	4 評価で首位、1 タスク$7.76、限定提供[1]
Claude Opus 5.5（max）	58	FrontierCode 2.0・HLE-Diamond で首位、Coding Agent Index 首位（62）、1 タスク$10.04[1,4]
GPT-6 Astra（max）	57	Terminal-Bench-Science・CritPt で首位、1 タスク$6.24[4]
GPT-6.1 Sol（max）	55	1 タスク$1.27（Opus 5.5 の約 8 分の 1 のコストで 4 点差）[4]
Gemini 4 Argon（high）	50	v5 公開時は AutomationBench-AA で首位。10 月 13 日に一般提供[3,4]
Claude Fable 5.1（max）	記載なし	上位 10 モデル中で最も高い 1 タスク$13.04、HLE-Diamond 60%[4]
Kimi K3.1 Pro	46	オープンウェイト最上位（重み未公開）。AA-Briefcase では Carbon より上[1,4]。重みは 11 月 4 日までに公開予定とされる[17]
DeepSeek V4.1 Flash 1017／GLM-5.3（max）	32	重み公開済みモデルの最上位[4]
Grok 4.7（xhigh）	未確認	AA-Briefcase 1638 で Carbon より上[1]

他の指標との整合性【事実】

- **Arena：**Text Arena で 1 位（1528。2 位 Argon は 1520）。Agent Arena は 2 万 1900 超の実セッションで 2 位（純改善+16.2%。首位 Opus 5.5 は+21.1%、3 位 Astra は+13.4%）で、タ

スク単価の中央値$1.80 は Opus 5.5 より 72%安い。WebDev Arena は 4 位（1768。3 位 Sonnet 5.5 との差は 2 点）[8,18]。

- **Vals：**Vals Index v2.2 で 2 位（70.36%。1 位 Sonnet 5.5 は 70.85%、3 位 Opus 5.5 は 70.18%）、1 テスト$10.51・約 48 分[8,19]。Finance Index と Legal Research で 1 位。Legal Research では Argon より引用ソースが約 45%多い一方、1 タスク平均 50 分超かかる[8]。推論トークンは Argon の 2.3 倍、出力トークンは 64%多い[8,20]。Legal Index は 3 位（75.83%）との報道もある[21]。
- 【分析】「総合・知識・科学ではトップ級、コーディングエージェントとコンピュータ操作では Anthropic が優位」という傾向は、各指標で一致している。

5. 反響・評価

- **英語圏メディア【事実】：**VentureBeat は「首位奪還、ただしアクセスは限定」と報じ、ハルシネーション率の悪化と導入価格期限の未定を指摘した[3]。The Decoder は、独立評価が Google の数値を裏づけたと評価した[8]。
- **開発者コミュニティ：**Hacker News では、発表しても使わせない方式への批判、命名の分かりにくさへの困惑、Terminal-Bench 不正 0%への好意的な反応が見られた[16]。Latent Space は各独立評価の結果を集約している[20]。
- **日本語メディア：**ITmedia、窓の杜など複数の媒体が、脆弱性対応の強化と防御者限定の提供を軸に報じた[22,23]。
- **中国語メディア：**新浪、鳳凰網、網易などが「谷歌重回世界第一（Google が世界一に返り咲き）」と大きく報じた。一部は Alphabet 決算直前という発表タイミングに注目し、コーディングエージェントでは Claude／GPT が依然リードしていると付け加えている[24–27]。
- **実使用での評判【事実】：**アクセスが限られているため、外部の実使用評価はほぼ存在しない。発表前に LMArena の匿名対戦で試した利用者の「3D 出力は改善したが遅い」といった報告がある程度である[28]。

6. 知財業務への影響予測

6.1 前提となる制度・実務環境【事実】

- **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 4 月）：**外部の生成 AI への秘密情報の入力は守秘義務（弁理士法 30 条）違反のおそれがあり、秘密保持規約のないプラットフォームへの入力は新規性喪失の可能性がある。クライアントの秘密情報を入力する際は学習の有無にかかわらず合意を取り、ファクトチェックは弁理士自身が行う、と整理している[29]。
- **特許庁：**「AI 技術の活用に向けたアクション・プラン（令和 7 年度改定版）」で生成 AI の審査業務への適用を項目として掲げた[30]（項目の本文は未検証）。2026 年 7 月の「JPO AI ビジョン」は二次情報でのみ確認した[31]。
- **USPTO：**AI 事前検索パイロット「ASAP!」（出願人に最大 10 件の先行技術を通知）は 2026 年 6 月 1 日に終了した[32]。実務家向けの AI 利用ガイダンス（2024 年 4 月）は現在も有効である[33]。
- **In re Mitchell（D2026-16）：**ペンシルベニア東部地区連邦地裁の特許訴訟（Magpul Industries v. Mission First Tactical）で、生成 AI で作成したクレーム解釈チャートに明細書・図面・審査経過への架空の引用が含まれていたとして、USPTO が代理人を公開戒告した。2 つ目の AI でレビューしていても検証義務を果たしたとはみなされず、地裁が制裁を科していなくても懲戒された。USPTO は、AI による引用誤りは制定法や判例などの外部ソースに限らず、出願書類や包袋にも及びうると注意喚起している[5]。存在しない引用は 47 件あり、そのうち 26 件は審査経過に帰属させられていた[34]。
- **EPO：**2026 年版審査ガイドライン（4 月 1 日施行）に AI 利用に関する項目が新設され、AI の利用有無にかかわらず書面の責任は当事者にあると整理された（法律事務所の解説で確認、原文は未検証）[35,36]。
- **Gemini API のデータ取扱い：**有償 API でも不正利用監視のためにプロンプトが一定期間保持されうる。保証されたゼロデータ保持が必要な場合、Google は Vertex AI（企業向け環境）を案内している[37]。
- **参考：**特許起草支援ツールのランダム化比較試験で、米国特許弁護士 133 名の起草スコアが 0.34〜0.38 標準偏差向上したとの報道があるが、作業論文の二次情報にとどまる[38]。

6.2 業務別の影響予測【予測】

表 4 業務別の影響予測

業務	影響度	Carbon 級モデルで変わること	主なリスク・留保
先行技術調査・無効資料調査	大	1M トークンで数十〜数百件の公報を一度に読ませ、クレーム要素ごとの対比表を作るエージェント処理が現実的になる。検索式の作成→精読→絞り込みを夜間バッチで回せる	DB の網羅性はモデルでは補えない。長文脈性能の根拠（GraphWalks）は Google 計測。日本語・中国語公報での精度は未検証
FTO（侵害予防調査）	中〜大	製品仕様書・図面とクレームを突き合わせる一次スクリーニングが効率化する	見落としが致命的。偽陰性（関連特許の取りこぼし）の評価が必須
明細書作成	中	発明提案書・図面・実験データからの下書き、実施例の展開	守秘義務・新規性喪失の問題。補正の根拠となる記載の欠落は後から直せない
中間処理（拒絶理由対応）	中〜大	引用文献と本願の対比、補正案・意見書の下書き。曖昧な点を推測で進めず確認する改善（Google 発表）は実務と相性が良い可能性	Mitchell 型の架空引用（段落・図番号の誤り）。原典照合の必須化
特許分析・IPL	大	大量の公報の分類・要約・技術マップ化をエージェントで実行できる。費用対効果が最も改善しやすい	集計値の誤りは目視で見つけにくい。サンプル検証の設計が必要
契約・ライセンス	中	条項比較、ロイヤルティ条件の抽出（Vals Legal Research で 1 位）	米国法中心のベンチ。日本法・日本語契約への適用性は未検証
知財戦略立案	小〜中	競合分析の材料作りは速くなる	判断そのものは代替できない。指標の首位は戦略の質を保証しない
審査実務への波及	中（中期）	各庁が AI による先行技術提示を広げるほど、出願前調査に求められる水準も上がる	審査側 AI が拾う文献を見越し、明細書の技術的裏づけを厚くする必要

6.3 観点別の評価【予測】

- **長文脈：**1M トークンと GraphWalks 96.2%は、包袋全体や大量の引用文献を扱ううえで有利である。ただし v5 は長文脈評価（AA-LCR）を外しており、59 点は長文脈性能の保証にならない。

- **多言語：**v5 も Google の発表も多言語の評価を含まない。日中韓の公報や拒絶理由通知での精度は、自社で評価するしかない。
- **マルチモーダル（図面）：**画像・動画の入力に対応し、図表の解釈や設計図の類似度判定の改善が報告されている[3]。しかし特許図面（符号の対応、断面図の読解）での実証はない。
- **エージェント能力：**「数日間走り続けた」というのは Google 社内の観察である[2]。AutomationBench-AA で首位、Agent Arena で 2 位という結果は、調査パイプラインの自動化には追い風となる。
- **コスト・速度：**1 タスク$7.76 は Opus 5.5 より安いが、GPT-6.1 Sol（$1.27）の約 6 倍である。初回応答まで約 100 秒、Legal Research では 1 タスク 50 分超かかるため、対話型の補助よりバッチ処理に向く。大量スクリーニングは安価なモデル、最終的な対比はフロンティアモデル、という使い分けが合理的である。

7. 推奨アクション（知財部門・特許事務所）

1. **今は乗り換えない：**Carbon は一般には使えず、API 仕様もデータ保持条件も公表されていない。実務は一般提供済みのモデル（Opus 5.5、GPT-6 Astra／6.1 Sol、Gemini 4 Argon など）で続け、Carbon は一般提供後に評価する。
2. **自社評価セットを今すぐ作る（最優先）：**過去案件から正解付きで 20〜50 件を選ぶ（無効資料調査で実際に効いた文献、拒絶理由通知と応答の組、日本語・中国語公報の対比表、図面符号の読解など）。指標は「主要文献の再現率」「架空引用・誤引用の件数」「1 件当たりの費用・所要時間」とし、AA の 59 点ではなくこの結果でモデルを選ぶ。
3. **ガバナンスを整える：**弁理士会ガイドラインに沿って、クライアントの同意取得と入力可能な情報の区分を定める（未公開の発明は原則として企業契約の環境に限る）。Google 系を使う場合は、ゼロデータ保持とデータ所在地を契約で確保できる企業向け環境の経由を前提とする。出願前の発明情報を消費者向けアプリに入力しない。
4. **検証義務を明文化する：**In re Mitchell を踏まえ、「AI による相互チェックは検証とみなさない」と所内規程に明記する。段落番号・図番号・頁行の引用は、人が原典で確認する。
5. **コストを管理する：**導入価格の終了日が確定していないため、試算は標準価格（$4／$20）で行う。トークン消費の多いモデルであることを前提に、1 件当たりの費用で予算を組む。

6. **動向を追う：**一般提供の時期、API 仕様、モデルカード・技術レポートの公開、競合各社の対抗モデル、Kimi K3.1 Pro の重み公開（11 月 4 日予定とされる）を注視する。

8. 留意事項

- **未検証事項：**AA の Carbon モデルページ（速度などの数値）、Fairwind Program のページ、Business Insider の原記事、特許庁アクション・プランの該当項目本文、JPO AI ビジョン本文、EPO ガイドライン原文、Google の評価手法 PDF は、本文を直接確認していない。
- **数値の食い違い：**WebDev Arena は 1768（Google）と 1769（The Decoder の一部記述）。Terminal-Bench 4.0 は 64%（AA の Antigravity CLI 評価）と 65.2%（Google の表）。Google の表で首位となった項目数は「17 行中 15」（VentureBeat）と「16 中 15」（The Decoder）。Argon の Coding Agent Index は 53（各記事本文）と 50（The Decoder の要点欄）で、本稿は 53 を採用した。いずれも評価条件や記載の違いによると考えられる。
- **自社公表値の性質：**サイバー系ベンチはガードレール無しでの測定であり、利用者が同じ結果を得られるとは限らない。社内指標（セキュアコード 2 倍超、Chrome 10 倍超など）は第三者が検証できない。
- **時点：**本稿は 2026 年 10 月 29 日時点の情報に基づく。一般提供や追加の独立評価により、評価が変わる可能性がある。

参考文献

[1] Artificial Analysis「Gemini 4 Carbon is the new #1 on the Artificial Analysis Intelligence Index v5」2026 年 10 月 28 日。 https://artificialanalysis.ai/articles/gemini-4-carbon-new-number-1

[2] Kavukcuoglu, K.（Google）「Gemini 4 Carbon: our next frontier model」Google The Keyword、2026 年 10 月 28 日。 https://blog.google/innovation-and-ai/models-and-research/gemini-models/introducing-gemini-4-carbon/

[3] Franzen, C.「Google unveils Gemini 4 Carbon, retaking the world's no. 1 spot on updated, harder benchmarks — but access is limited」VentureBeat、2026 年 10 月 28 日。 https://venturebeat.com/technology/google-unveils-gemini-4-carbon-retaking-the-worlds-no-1-spot-on-updated-harder-benchmarks-but-access-is-limited

[4] Artificial Analysis「Announcing Artificial Analysis Intelligence Index v5」2026 年 10 月 27 日。 https://artificialanalysis.ai/articles/artificial-analysis-intelligence-index-v5

[5] Nikolaidis, A.「USPTO Issues Its First AI-Predicated Discipline Order Involving Hallucinated Cites to Intrinsic Record」IPWatchdog、2026 年 8 月 28 日。 https://ipwatchdog.com/2026/08/28/uspto-issues-its-first-ai-predicated-discipline-order-involving-hallucinated-cites-to-intrinsic-record/

[6] Artificial Analysis「Gemini 4 Carbon」モデルページ。 https://artificialanalysis.ai/models/gemini-4-carbon **【未検証】**

[7] R&D World「Google's Gemini 4 Carbon takes benchmark lead with access limited to cyber defenders」2026 年 10 月。 https://www.rdworldonline.com/googles-gemini-4-carbon-takes-benchmark-lead-with-access-limited-to-cyber-defenders/ ※

[8] Bastian, M.「Google Deepmind releases Gemini 4 Carbon just two weeks after Argon with gains across the board」The Decoder、2026 年 10 月 28 日。 https://the-decoder.com/google-deepmind-releases-gemini-4-carbon-just-two-weeks-after-argon-with-gains-across-the-board/

[9] SaaS City「Gemini 4 Carbon: Benchmarks, Price and When You Can Actually Use It」2026 年 10 月 28 日。 https://saascity.io/blog/gemini-4-carbon-benchmarks-pricing-availability-2026 ※

[10] Kingy AI「Gemini 4 Carbon: Specs, Benchmarks, Pricing and Availability」2026 年 10 月。 https://kingy.ai/blog/gemini-4-carbon/ ※

[11] Fello AI「Gemini 4 Carbon: Benchmarks, Price and Release Date」2026 年 10 月 28 日。 https://felloai.com/gemini-4-carbon/ ※

[12] IT 之家「谷歌 Gemini 4“Argon”模型即将发布，消息称内部测试“Carbon”新版本」2026 年 10 月 10 日（Business Insider 報道の紹介。原記事は未確認）。 https://ithome.com/1/011/162.htm ※

[13] X 投稿（@astropol0）、2026 年 10 月。 https://x.com/astropol0/status/2115032943253221608 **【未検証】**

[14] Lambert, N.「Artificial Analysis v5 and the benchmarkers」Interconnects、2026 年 10 月。 https://www.interconnects.ai/p/artificial-analysis-v5-and-the-benchmarkers ※

[15] Hacker News、Artificial Analysis Intelligence Index v5 に関するスレッド、2026 年 10 月。 https://news.ycombinator.com/item?id=50268868 ※

[16] Hacker News、Gemini 4 Carbon に関するスレッド、2026 年 10 月。 https://news.ycombinator.com/item?id=50276792 ※

[17] Yotta Labs「Kimi K3.1 Pro: specs, price, open weights, benchmarks」2026 年 10 月。 https://www.yottalabs.ai/post/kimi-k3-1-pro-specs-price-open-weights-benchmarks ※

[18] Arena.ai（@arena）X 投稿、2026 年 10 月 28 日。 https://x.com/arena/status/2115495002998751262 ※

[19] Vals AI「Vals Index」（2026 年 10 月 27 日時点）。 https://www.vals.ai/benchmarks/vals_index ※

[20] Latent Space「[AINews] Gemini 4 Carbon tops leaderboards, launches for cyber defenders」2026 年 10 月 28 日。 https://latent.space/p/ainews-gemini-4-carbon/ ※

[21] Intelligent Living「Gemini 4 Carbon Takes the AI Crown at Half the Price」2026 年 10 月。 https://www.intelligentliving.co/gemini-4-carbon-ai-crown-price/ ※

[22] ITmedia NEWS、Gemini 4 Carbon 発表に関する記事、2026 年 10 月 29 日。 https://www.itmedia.co.jp/news/article/2610/29/2000002690/ ※

[23] 窓の杜（Impress）、Gemini 4 Carbon 発表に関する記事、2026 年 10 月。 https://forest.watch.impress.co.jp/docs/news/2152085.html ※

[24] 新浪財経、Gemini 4 Carbon 発表に関する記事、2026 年 10 月 29 日。 https://finance.sina.com.cn/tech/roll/2026-10-29/doc-iniwvwrf8972543.shtml ※

[25] 鳳凰網科技、Gemini 4 Carbon 発表に関する記事、2026 年 10 月。 https://tech.ifeng.com/c/8xaMcSLMa0n ※

[26] 網易、Gemini 4 Carbon 発表に関する記事、2026 年 10 月。 https://www.163.com/dy/article/LAD7F0PV05399DAP.html ※

[27] 投資界、Gemini 4 Carbon 発表に関する記事、2026 年 10 月。 https://news.pedaily.cn/202610/571031.shtml ※

[28] TestingCatalog「Gemini 4 Carbon early test results are in ahead of public release」2026 年 10 月。 https://testingcatalog.com/gemini-4-carbon-early-test-results ※

[29] 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 4 月。 https://www.jpaa.or.jp/cms/wp-content/uploads/2025/04/AIservices-guideline.pdf ※

[30] 特許庁「AI 技術の活用に向けたアクション・プラン（令和 7 年度改定版）」令和 7 年 6 月 25 日（該当項目の本文は未検証）。 https://www.jpo.go.jp/system/laws/sesaku/ai_action_plan/ai_action_plan-fy2025.html ※

[31] Japan Security Summit、JPO AI ビジョンに関する記事、2026 年 7 月。 https://japansecuritysummit.org/2026/07/14740/ **【未検証】**

[32] Patrick IP Law「USPTO AI prior art search tools are improving; applicants should prepare accordingly」。 https://blog.patrickiplaw.com/uspto-ai-prior-art-search-tools-are-improving-applicants-should-prepare-accordingly ※

[33] USPTO「Guidance on use of AI tools in practice before the USPTO」。 https://www.uspto.gov/learning-and-resources/uspto-videos/guidance-use-ai-tools-practice-uspto ※

[34] Kline, G.「In re Mitchell: Why the Second AI Didn't Save Him」IPWatchdog、2026 年 9 月 29 日。 https://ipwatchdog.com/2026/09/29/mitchell-why-second-ai-didnt-save-him/

[35] Wilson Gunn「Five key changes to the EPO Guidelines」。 https://www.wilsongunn.com/resource/news/five-key-changes-to-the-epo-guidelines ※

[36] Venner Shipley「EPO publishes preview of the 2026 European Patent Office Guidelines」。 https://www.vennershipley.com/insights-events/epo-publishes-preview-of-the-2026-european-patent-office-guidelines/ ※

[37] Google AI for Developers「Zero data retention」（Gemini API ドキュメント）。 https://ai.google.dev/gemini-api/docs/zdr ※

[38] TokenPost、特許起草支援ツールのランダム化比較試験に関する記事（二次情報）。 https://www.tokenpost.com/news/technology/29543 ※