

GLM-5.2調査報告

エグゼクティブサマリー

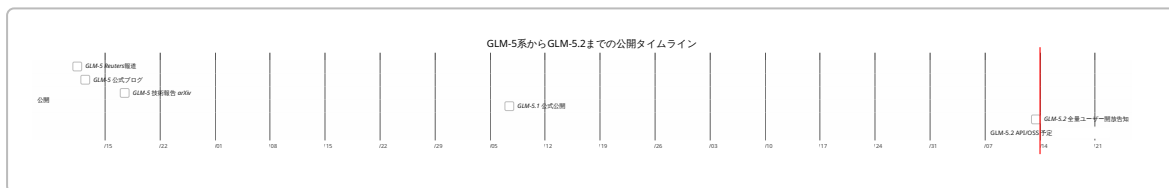
2026年6月14日時点で、GLM-5.2は実在し、Z.AIのCoding Plan上ではすでに利用可能です。公式ドキュメントでは、GLM-5.2はLite/Pro/Maxを含む全プランで利用でき、`glm-5.2[1m]` という指定で1Mトークン文脈を有効化できること、さらに対応ツールではコンテキスト圧縮ウィンドウを1,000,000に設定するよう明示されています。一方で、一般向け公開APIとオープンソース重みは「来週」提供予定とされており、2026年6月14日時点ではまだ公開済みの一次資料を確認できません。ライセンスについては、6月13日の公式告知を引用する複数の報道がMITでの公開予定を伝えています。 ¹

したがって、GLM-5.2は「製品としては出ているが、研究・ガバナンス面の公開情報はまだ追いついていない」状態だと見るのが妥当です。現時点でアーキテクチャ、パラメータ数、学習データ増分、最大出力長、モダリティ、安全性レポート、一般API価格表は未公表です。技術的に確認可能な基盤は、現状ではGLM-5/GLM-5.1系で、こちらは744B total / 40B activeのMoE系、28.5Tトークン学習、DSA、非同期RL基盤 Slimeまで公開されています。 ²

知財実務への適用という観点では、GLM-5.2は最終判断を自律委任する法務モデルというより、長文脈・証拠整理・構造化出力・ツール連携を担当する“オーケストレーション層”として使うのが最も合理的です。とくに、先行技術検索、クレーム分解、FTO一次スクリーニング、IP DD、契約レビューのような「大量資料を保持しつつ、証拠を引きながら比較・抽出する」作業と相性が良い一方、法的判断、クレームの最終ドラフティング、侵害/非侵害の結論は人間の弁理士・弁護士のサインオフが前提です。これはGLM-5.2自体の欠点というより、現時点で安全性カードや法務特化評価が十分に公開されていないためです。 ³

確認済み仕様と公開状況

まず、確認できる事実を切り分けると、GLM-5.2はCoding Plan上での最新モデルとして文書化済みであり、一般API・OSSの完全公開はこれからです。Z.AI公式ドキュメントは、Claude Codeなどの対応ツールでGLM-5.2へ切り替える方法を案内しており、さらに中国メディアは、6月13日の智譜公式告知として「当日17:21から全量開放、APIは来週、OSSも来週、MIT」を伝えています。 ⁴



このタイムラインで重要なのは、GLM-5.2だけが現時点で「製品ドキュメント先行・研究資料未追隨」という点です。GLM-5は技術報告とOSSがそろっていましたが、GLM-5.2については、少なくともレビューした一次資料群ではまだその段階に達していません。 ⁵

論点	2026-06-14時点で確認できること	まだ未確認・保留
リリース	6月13日に全量ユーザー開放告知。Coding Plan利用者向けに展開。	一般公開APIの正式公開日時、OSS公開ページ。

論点	2026-06-14時点で確認できること	まだ未確認・保留
提供形態	Coding Planの 公式対応ツール内 で利用。	独自社内アプリ向け一般提供の範囲。
コンテキスト長	1M 。 <code>glm-5.2[1m]</code> と <code>1000000</code> の圧縮ウィンドウ設定が明記。	1Mが 完全ネイティブ 注意か、圧縮/外挿/メモリ併用か。
API	専用Coding APIは存在するが、 Coding Plan特典は公式対応ツール用途に限定 。一般APIは来週予定。	一般APIのエンドポイント詳細、レート制限、価格表。
OSS/ライセンス	6月13日の公式告知ベース報道では、 来週MITで公開予定 。	実際のモデルカード、配布先、重みの利用条件全文。
価格	Lite \$18/月、Pro \$72/月、Max \$160/月。GLM-5.2/5-Turboはピーク3倍・オフピーク2倍、期間限定でオフピーク1倍。	トークン単価ベースのGLM-5.2価格。
併存モデル	docs上は GLM-5.2、GLM-5-Turbo、GLM-4.7、GLM-4.5-Air等が併存。	GLM-5.2の軽量版・量子化版・FP8版の公開計画。

注記として、**既存の GLM-5/GLM-5.1 系では、GitHub repoのコードライセンスが Apache-2.0、Hugging Face上の重みライセンスが MIT**という切り分けになっています。GLM-5.2でも同じ整理になる可能性はありますが、現時点ではまだ確認がありません。 [6](#)

技術アーキテクチャと革新点

最重要の注意点として、**GLM-5.2固有の技術報告はまだ見当たりません**。そのため、ここで「確認済み」と言えるのは、**GLM-5/GLM-5.1までで公開された土台技術と、GLM-5.2 docsで新たに露出したプロダクト仕様**の二層です。5.2の内部構造を、5.1と同一だと断定するのは避けるべきです。 [7](#)

GLM-5系の公開済み基盤としては、**355B/32B active から 744B/40B active へのスケールアップ、23Tから 28.5Tトークンへの学習データ拡張、DeepSeek Sparse Attention の統合、そして長期タスク学習を回すための非同期RL基盤 “slime”**が明示されています。GLM-5 repoはこれを簡潔にまとめており、技術報告の要約も同じ方向です。 [8](#)

学習データについては、公開範囲内でも比較的情報量があります。技術報告は、**27T規模の高品質 Web+Codeコーパス**を中心に、一般Web、オープンソースコードホスティング、Software Heritage由来コード、教育サイト・教科書・技術文書・学術PDF由来の数理/科学データを使い、そこに**9T相当の code & reasoning**を加えて**総計28.5T**にしたと説明しています。さらに、ソフトウェア工学特化データとして、**コード本体、commit diff、GitHub issue、PR、関連ソースファイル**を含むデータを構築し、**約1,000万件の issue-PR ペア**を残し、**約160B unique tokens**を得たとしています。 [9](#)

長文脈学習も段階的です。技術報告は、**32K・128K・200K**へと順次伸ばす mid-training を採り、**長いコード/推論データや長文書・合成agentトラジェクトリ**を追加したと述べています。したがって、GLM-5.2の1M文脈は「突然現れた機能」ではなく、**5系で長文脈と長期タスクを継続的に強化してきた延長線上**にあります。もっとも、**1Mに到達した方式そのものは現時点で未公開**です。 [10](#)

GLM-5.1で追加的に確認できるのは、**200K文脈・128K最大出力、最大8時間の長期自律タスク、Agentic Coding向け強化**です。Z.AI公式 docs は、GLM-5.1が単一タスクに対して **planning→execution→iterative**

optimization を長時間継続できるとし、ベクトルDB最適化やKernelBench Level 3最適化の事例を出しています。これは、5.2の「agentic coding」 定位を理解する上での直近の公開材料です。 ¹¹

GLM-5.2で新たに確認できる技術的特徴は、現状では主にプロダクト面です。docs上では、`glm-5.2[1m]` による1M文脈、`/effort` 切替で **high** と **max** 系の思考強度マッピング、そして **Claude Code / Cursor / Cline / OpenCode / Roo Code / Goose** など多数の外部coding agentツールへの公式対応が明示されています。とくに、GLM Coding PlanがOpenAI互換設定で接続される設計は、IP用途でもツールチェーン統合しやすいという含意があります。 ¹²

ただし、1M文脈については慎重であるべきです。docsは `CLAUDE_CODE_AUTO_COMPACT_WINDOW=1000000` という圧縮前提の設定を要求しており、これは少なくともツール側の文脈圧縮・整理が効いていることを示唆します。したがって、現時点では「GLM-5.2は1M完全ネイティブ注意を持つ」と断言せず、“**1Mを実用化したプロダクト仕様は確認できるが、アーキテクチャ上の実現方式は未公開**”と書くのが一番厳密です。 ¹³

公式ベンチマークと独立評価

GLM-5.2自身の公開ベンチ表は、2026-06-14時点では未確認です。そのため、公式実績として参照できるのは、GLM-5 論文/README と GLM-5.1 docs です。ここを5.2の“周辺証拠”として扱い、5.2固有成績とは混同しないのが重要です。 ¹⁴

ベンチマーク	GLM-5	GLM-5.1	読み方
HLE	30.5	—	学術推論の基礎性能。
HLE with tools	50.4	—	ツール利用込みの難問処理。
HMMT Feb	97.9	—	数学競技系。
LongBench v2	64.5	—	長文脈理解。
SWE-Bench Verified	77.8	—	実ソフトウェア修正能力。
SWE-Bench Multilingual	73.3	—	多言語コード修正。
Terminal-Bench 2.0	56.2	公式READMEでは「GLM-5を大幅リード」とのみ明記	実端末での逐次タスク。
BrowseComp	62.0 / 75.9	—	Web探索と情報統合。
MCP-Atlas	67.8	—	MCP的なツール統合。
τ^2 -Bench	89.7	—	複雑な多ツール実行。
Vending Bench 2	\\$4,432	—	長期運用型エージェント。
SWE-Bench Pro	—	58.4	GLM-5.1 docsの主張ではSOTA。
KernelBench L3最適化	—	3.6× 幾何平均速度向上	長期最適化能力。

ベンチマーク	GLM-5	GLM-5.1	読み方
ベクトルDB最適化	—	初期版比 6.9×	655反復の長期改善。

この表から読み取れるのは、**GLM-5はすでに“コード生成”から“工程実行”へ振れたモデル**であり、5.1ではさらに**長期反復の持続性**を押し出していることです。5.2が本当にこの系譜を継ぐなら、IP用途でも単発要約より**多段比較・変数管理・証拠整理**に価値が出やすいはずですが、ただし、その最後の推論は**5.2に関する分析的推測**であって、まだ公式ベンチで裏づけられていません。 ¹⁵

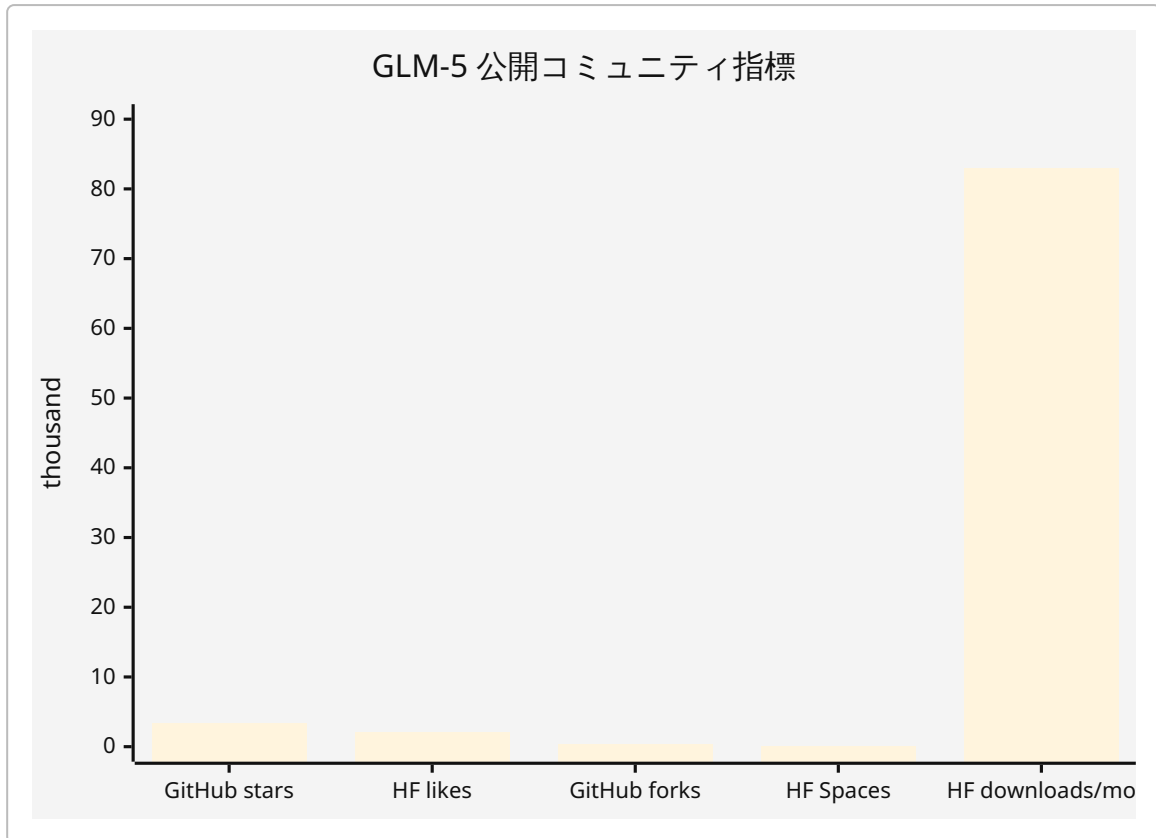
独立評価では、**Artificial Analysis**がGLM-5を**open-weights最上位**として扱い、**Intelligence Index 50、Agentic Index 63**、後者は**open-weights首位かつ全体3位**と報じました。**LiveBench**ではGLM 5が**68.85**。一方、**LM Arena**の2026年6月10日時点スナップショットでは、GLM-5は**Text全体36位 (1457)**、**Code/WebDev全体32位 (1435)**で、**GLM-5.1はCode/WebDev 10位 (1532)**、さらに**Software & IT Services**カテゴリでも**10位 (1516)**まで改善しています。これは、**公式ベンチでは非常に強い一方、公開人気投票型・実利用型リーダーボードでは5.1がようやくトップ層に食い込んだ**という読み方ができます。 ¹⁶

幻覚・頑健性・安全性については、ここが最大の空白です。VentureBeatは、z.aiの資料をもとに“**record low hallucination rate**”と要約していますが、これは**独立再測定ではなくベンダー資料の紹介**です。レビューした範囲では、**GLM-5.2の公開システムカードも、5.2固有の外部安全性監査も見当たりません**。したがって、知財導入時には「強いかわいかな」よりも、**安全性がまだ十分に文書化されていない**という前提で設計すべきです。 ¹⁷

コミュニティ評価と報道動向

中国市場での反応はかなり強いです。ReutersはGLM-5発売時に、**ZhipuがGLM coding planの価格を少なくとも30%引き上げるほど需要が急増した**と報じました。別のReuters記事では、2025年の**API収益が292%増、利用量が400%増、国内チップ最適化が進んだ**ことも報じています。中国メディアでも、GLM-5を**中国AIの“実行力”を示す転機**として扱う論調が目立ちました。 ¹⁸

グローバル側の受け止めも、概ね「**中国発のopen-weight coding/agentモデルが本格的にトップ層へ接近した**」というものです。ReutersはGLM-5を**Claude Opus 4.5やGemini 3 Proに競う中国の新フラッグシップ**として報じ、Artificial Analysisは**open-weights首位**と評価しました。LM Arenaの changelog を見ると、**GLM-5は2月11/12日にText/Code leaderboardsへ、GLM-5.1は4月7/9日に追加**されており、公開コミュニティ内での追跡対象としても定着しています。 ¹⁹

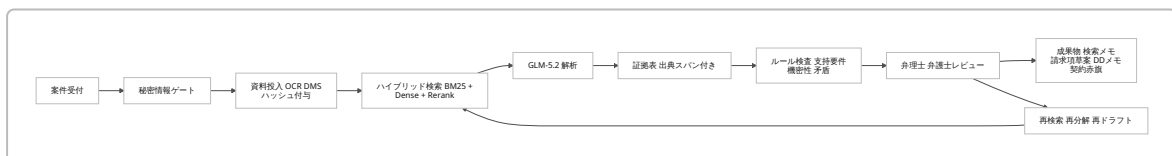


数値で見ると、GitHub repo は約3.4k stars、375 forks、32 issues、41 commits、Hugging Face の `zai-org/GLM-5` は2.09k likes、82,925 downloads/month、100 Spaces、38 finetunes、20 quantizationsです。これは「研究コミュニティの熱狂」だけでなく、運用・蒸留・量子化・周辺ツール化がすでに始まっていることを示します。知財用途で重要なのは、単に強いモデルというより、ローカル運用・派生モデル・検証ツールが育ちやすい生態系があることです。 ²⁰

他方で、評判には温度差もあります。公式ベンチや一部第三者指標ではGLM-5/5.1は非常に強い一方、安全性資料、正式な5.2ベンチ表、法務特化評価、長文脈の再現性検証はまだ薄く、これがエンタープライズ導入ではボトルネックになります。要するに、「開発者人気は高いが、知財部門がそのまま本番投入できるほど監査資料がそろっているわけではない」というのが、現時点でのもっともバランスの取れた評価です。 ²¹

知財実務への適用設計

GLM-5.2を知財で使うなら、“法律家の代替”ではなく、“証拠付きの下準備を高速に回す長文脈エンジン”として設計するのがよいです。根拠は三つあります。第一に、1M文脈が確認できること。第二に、Function Calling / Structured Output / Context Caching / MCPがGLM-5.1系で公式対応していること。第三に、現時点のCoding Planが公式対応ツール前提であり、自由なSDK統合には制約があることです。したがって、いま本番設計するなら、サポートされたエージェント環境での検証か、GLM-5/5.1 open weightsのローカル実装を先行させるのが安全です。 ²²



ユースケース	推奨ワークフロー	プロンプトの核	推奨ツールチェーン	主な限界と緩和策
特許ドラフティング	発明開示→先行技術サマリー→クレーム候補→支持表→人手修正	「以下の発明開示と引用文献だけを根拠に、独立請求項3案と、各要件の支持箇所表をJSONで作成せよ。根拠のない機能語は追加しない。」	社内DMS、特許全文DB、BM25+embedding再ランキング、JSON schema、クレーム支持表	過広クレーム、支持欠如、発明の一般化しすぎが最大リスク。対策は、 全要件に証拠スパン必須 、新規語の追加禁止、最終は弁理士レビュー。
先行技術検索	クレーム分解→同義語/上位概念/否定語生成→多言語検索→ファミリー統合→新規性マップ	「請求項を構成要件へ分解し、日本語・英語・中国語の検索語、除外語、近傍技術語、引用すべき非特許文献候補を作れ。」	USPTO/EPO/J-PlatPat系コーパス、NPL検索、family dedup、reranker、クラストリング	Recall不足 が本質的課題。対策は、 Recall@kを定例計測 し、機械抽出結果を検索担当者がtop-k再確認。
クレーム分析	クレーム要件化→対象製品資料投入→要件別充足/未充足/不明→証拠表	「各要件について‘充足/未充足/不明’を付し、各判断に引用文、ページ、理由を添えよ。不明を強制的に0にするな。」	1M dossier pack、出典スパン抽出、表形式出力、比較UI	誤った断定 が危険。対策は、“ 不明 ”を 正答にする設計 、引用のない判断を自動棄却。
FTO一次分析	製品仕様→権利有効性/法的状態でフィルタ→主要独立請求項チャート→リスク帯分け	「失効・取下げ・未維持の案件は除外し、存続・対象国一致・クレーム入手済み案件のみで一次FTO表を作れ。」	法的状態DB、family normalizer、claim chart UI、freshness timestamp	法的状態の鮮度と均等論/クレーム解釈 はLLMだけでは危険。対策は、 日付印付き法的状態確認と弁護士サインオフ必須 。
IP デューデリジェンス	データルーム投入→権利者/譲渡/共同研究/係争/OEM/OSSを抽出→赤旗一覧→不足資料要求	「チェーン・オブ・タイトル、職務発明、ライセンス負担、共同所有、係争、OSS汚染の赤旗を優先度順に列挙し、根拠文書IDを付けよ。」	OCR、entity resolution、graph DB、チェックリスト、監査ログ	欠落資料とOCR欠陥 が支配的。対策は、 文書完備性チェック、ページハッシュ保存 、欠落時は必ず“ 要追加提出 ”を返す。

ユースケース	推奨ワークフロー	プロンプトの核	推奨ツールチェーン	主な限界と緩和策
契約レビュー	条項分割→IP ownership / background-foreground IP / 守秘 / indemnity / LoL / 監査権抽出→プレイブック比較	「IP ownership、license-back、confidentiality、indemnity、limitation of liability、assignment 禁止条項を抽出し、当社プレイブックとの差分だけを危険度付きで示せ。」	OCR、条項セグメンテーション、構造化出力、レビュー承認フロー	条項の交差参照漏れが多い。対策は、 cross-reference validator と、人手での最終赤入れ。

この表は、GLM-5.2の確認済み能力から逆算した推奨設計です。現時点で「Z.AIが法務/知財でこのワークフローを公式推奨している」という意味ではありません。むしろ厳密には、**長文脈・構造化出力・ツール連携・長期タスク維持**という公開Capabilityを、知財に移植した分析提案です。 ²³

運用上のガードレールは、**精度・機密・プロベナンス**の三本柱で十分に設計すべきです。精度面では、**引用スパンのない結論を無効化し、“不明”の選択肢を残す**こと。機密面では、**GLM-5.2のOSSがまだ出ていない以上、最高機密案件では現行のGLM-5/5.1 open weights をローカル配備する**ほうが安全です。プロベナンス面では、**文書ID・ページ・段落・ハッシュ**を成果物と一体管理し、検索メモ・ドラフト・DDメモのすべてに**根拠逆参照**を持たせるべきです。さらに、Coding Planは**公式対応ツール外でのSDK利用を許容していない**ので、社内案件管理システムへ直接つなぐ構成は、一般APIまたはOSS公開後に見直すのが妥当です。 ²⁴

比較表と評価設計

まず、現時点でのGLM-5.2と主要 open / closed モデルの比較を、知財導入しやすさの観点で並べると次のようになります。ここでは、**2026-06-14時点で確認できる範囲だけ**を書き、未公表は未公表のまま残しています。 ²⁵

モデル	コンテキスト長	ライセンス	API可用性	コスト	強み	弱み
GLM-5.2	1M	MIT予定だが未公開	Coding Plan内では可。 一般APIは来週予定	月額Lite/Pro/Max = \$18/\$72/\$160。 トークン単価は未公表	1M文脈、agentic coding 定位、対応ツールが広い	研究報告・安全性資料・公開ベンチ表が未整備
GLM-5.1	200K	MIT (重み)。repo コードは Apache-2.0	あり	\$1.4 / \$4.4	open-weight で高性能、ローカル配備可能、長期タスク強い	5.2より文脈短い。5.2相当の最新仕様は未確認

モデル	コンテキスト長	ライセンス	API可用性	コスト	強み	弱み
Llama 3.1 405B	128K	Llama 3.1 Community License	Llama API familyあり	reviewed sourceでは 3.1固定価格未確認 。重みDL可	エコシステムが圧倒的、カスタマイズしやすい	標準ライセンスがOSI系ではなく、最新 agentic codingでは相対的に古い
Mistral Medium 3.5	256K	Modified MIT	あり	\$1.5 / \$7.5	open weights + API、agentic/coding最適化、256K	1Mではない。価格は小型openモデルより高い
GPT-4o	128K	Proprietary	あり	\$2.5 / \$10	API成熟度、マルチモーダル、周辺ツールが豊富	クローズド、ローカル配備不可、文脈は1M級でない
Claude Opus 4.6	1M beta	Proprietary	あり	\$5 / \$25 、200K超は premium pricing	長期 agentic codingの参照点、1M対応	高価、クローズド

表の使い分けとしては、今日すぐに機密性の高い知財業務を回すなら **GLM-5.1 self-hosted**、**1M文脈の将来投資**として**GLM-5.2**を追跡、閉源でもよいなら **Claude Opus 4.6 / GPT-4o**を比較対象、**open weights と API のバランス**なら **Mistral Medium 3.5**、という整理が実務的です。Llama 3.1はエコシステム面で依然強いですが、**2026年の agentic coding 最前線**という意味では比較軸が一段ずれます。 26

知財向けの評価設計は、**公開ベンチマークだけでは足りません**。理由は単純で、先行技術検索やFTOの価値は、一般コーディング課題の pass@1ではなく、**証拠の取りこぼしがどれだけ少ないか**で決まるからです。したがって、公開データセットはサニティチェックにとどめ、**社内過去案件・外部法律意見・拒絶理由通知・検索報告を起点にした金標準**を作るべきです。 27

タスク	推奨データセット構成	自動評価指標	人手評価ゲート
特許ドラフティング	社内過去発明開示 50-100件 + 公開サニティ用に BigPatent など	支持表カバー率、未支持要件率、新規語混入率、引用スパン欠落率	弁理士が「補正可能」「再設計必要」を判定

タスク	推奨データセット構成	自動評価指標	人手評価ゲート
先行技術 検索	過去検索報告・拒絶理由 通知・IDS から relevance set 作成	Recall@10/20、NDCG@20、family dedup precision、novelty-hit@k	検索担当者が top-k を再査読
クレーム 分析	社内/外部の既存 claim chart	要件抽出F1、証拠スパン exact match、不明ラベル適正率	弁護士/弁理士が“断 定しすぎ”を監査
FTO	製品仕様 × 国 × 存続特 許の過去案件	active-claim recall、status freshness、false-clearance rate	国別 counsel が clearance 妥当性確 認
IP DD	過去 data room ・ 投資案 件 ・ M&A案件	red-flag recall、chain-of-title completeness、矛盾検出率	DD責任者が赤旗の 優先順位を確定
契約レ ビュー	自社ひな形、相手方ひな 形、CUAD系スキーマの拡 張	clause extraction F1、deviation precision/recall、cross-reference resolution	法務担当が redline の採否を決定

プロンプト評価の設計も、単発回答の品質ではなく、最終成果物の監査可能性を測るべきです。おすすめは、各出力に対して「根拠文書ID」「引用スパン」「確信度」「不足情報」を必須欄にし、これが埋まらないと回答を失敗扱いにすることです。知財業務で重要なのは“それっぽさ”ではなく、後から誰が見ても追試できることです。 28

最終的な導入判断としては、いま本番採用するなら GLM-5.2 を全面採用するより、GLM-5.1 の self-host / API と比較検証し、5.2 の一般API・OSS・ベンチ・ガバナンス資料が揃うのを待って昇格判定するのが最も堅実です。逆に、非機密の大規模ドキュメント整理・クレームマッピング・検索語生成のような周辺工程なら、GLM-5.2の1M文脈を早期検証する価値は十分あります。 29

🔗navlist🔗最近の関連報道🔗turn39news39,turn39news38,turn49news32🔗

1 4 12 13 23 <https://docs.z.ai/devpack/latest-model>

<https://docs.z.ai/devpack/latest-model>

2 7 8 20 <https://github.com/zai-org/GLM-5>

<https://github.com/zai-org/GLM-5>

3 <https://docs.z.ai/devpack/tool/others>

<https://docs.z.ai/devpack/tool/others>

5 <https://z.ai/blog/glm-5>

<https://z.ai/blog/glm-5>

6 28 <https://docs.z.ai/devpack/quick-start>

<https://docs.z.ai/devpack/quick-start>

9 10 15 <https://arxiv.org/pdf/2602.15763>

<https://arxiv.org/pdf/2602.15763>

11 26 <https://docs.z.ai/guides/llm/glm-5.1>

<https://docs.z.ai/guides/llm/glm-5.1>

14 21 25 29 [https://www.sohu.com/a/1036104418_114760?
scm=10001.325_13-325_13.0.0-0-0-0.5_1334](https://www.sohu.com/a/1036104418_114760?scm=10001.325_13-325_13.0.0-0-0-0.5_1334)
https://www.sohu.com/a/1036104418_114760?scm=10001.325_13-325_13.0.0-0-0-0.5_1334

16 <https://artificialanalysis.ai/articles/glm-5-everything-you-need-to-know>
<https://artificialanalysis.ai/articles/glm-5-everything-you-need-to-know>

17 [https://venturebeat.com/technology/z-ais-open-source-glm-5-achieves-record-low-hallucination-
rate-and-leverages](https://venturebeat.com/technology/z-ais-open-source-glm-5-achieves-record-low-hallucination-rate-and-leverages)
<https://venturebeat.com/technology/z-ais-open-source-glm-5-achieves-record-low-hallucination-rate-and-leverages>

18 [https://www.reuters.com/technology/chinese-ai-startup-zhipu-hikes-prices-coding-plan-demand-
rises-2026-02-12/](https://www.reuters.com/technology/chinese-ai-startup-zhipu-hikes-prices-coding-plan-demand-rises-2026-02-12/)
<https://www.reuters.com/technology/chinese-ai-startup-zhipu-hikes-prices-coding-plan-demand-rises-2026-02-12/>

19 [https://www.reuters.com/technology/chinas-ai-startup-zhipu-releases-new-flagship-model-
glm-5-2026-02-11/](https://www.reuters.com/technology/chinas-ai-startup-zhipu-releases-new-flagship-model-glm-5-2026-02-11/)
<https://www.reuters.com/technology/chinas-ai-startup-zhipu-releases-new-flagship-model-glm-5-2026-02-11/>

22 24 <https://docs.z.ai/devpack/overview>
<https://docs.z.ai/devpack/overview>

27 <https://livebench.ai/>
<https://livebench.ai/>