

初の“長考”できる国産フルスクラッチLLM「PLaMo 3.0 Prime」 深掘り分析レポート

エグゼクティブサマリ

株式会社Preferred Networks ¹ (PFN) は2026年3月19日、国産生成AI基盤モデルPLaMoの最新フラッグシップとして「PLaMo 3.0 Prime β版」を公開し、商用版提供に先立つモニター企業募集を開始しました。

² 同社は本モデルを「Reasoning (推論) 能力を中核に据えたモデル」と位置づけ、複数条件を整理して段階的に結論へ導く「信頼できる思考プロセス」を特徴として明示しています。 ³

技術的に公表されている“長考 (long-form reasoning)”の実体は、(1) 推論モデルとしての事後学習 (SFT/DPO/RL) を明示し、回答だけでなく**思考過程にも損失をかける**学習へ変更したこと、(2) 推論時に「思考過程を含む出力」を生成できること、(3) 推論トークン上限 (max_reasoning_tokens) を設定して評価していること、の3点に集約できます。 ⁴ 一方で、アーキテクチャ詳細 (Dense/MoE、層数、総パラメータ数など) や事前学習データ規模、言語比率は**未公開**です。 ⁵

性能面では、PFNが公開したベンチマーク図 (11指標) において、指示追従 (IFBench/JFBench) や日英対話 (MTBench/Japanese MTBench) で、比較対象のQwen3-235B系やgpt-oss-120b (reasoning effort=medium) と同等～一部上回るスコアを示しています。 ⁶ 反対に、数学推論 (AIME 2024) や複数ターン複数ツール選択 (BFCLの最難条件)、STEM難問 (GPQA-Diamond) では差が残ることも同じ図と本文で明確です。 ⁷

提供形態・条件について、β版は「無償利用を前提にモニター企業」を募り、実運用でのトラフィック量や応答性能を検証しつつ、安定性・処理速度を改善した上で**2026年6月中旬に商用版として正式提供予定**とされています。 ² ただし、βモニターの具体条件 (商用利用可否、SLA、データ保持、秘密保持など) は公開情報だけでは**未公開**です。 ²

なお、参照指定のITmedia AI+ ⁸ 記事は、本環境上の取得制約により本文の完全取得ができなかったため、検索抜粋・外部要約表示で確認できる範囲をPFN一次情報で突合しました (一次情報の結論を優先)。 ⁹

技術分析

PFN技術ブログによれば、PLaMo 3.0 Prime β版は「アーキテクチャを一新し事前学習からゼロベースで再構築」したフルスクラッチ・フラッグシップであり、βとして提供されます。 ¹⁰ ただし、この“新アーキテクチャ”がDenseかMoEか、層数や総パラメータ数がいくつか等は**未公開**です (少なくとも当該プレス・ブログには記載なし)。 ²

一方、PFNが国立研究開発法人情報通信研究機構 ¹¹ と進める次世代「PLaMo 3」系 (8B/31B事前学習検証) では、PLaMo 2で採用したSamba系から転換し、**Attention+Sliding Window Attention** (Google Gemma 3に近い) へ変更したこと、さらに長系列でKVキャッシュ消費を抑える比較 (例: 系列長128KでPLaMo 3 31Bは2.2GiB) を示しています。 ¹² PLaMo 3.0 Prime β版が同一設計かは明示されていないため断定はできませんが、「PLaMo 3.x Prime系統」として継続開発する起点モデルであることから、少なくとも“PLaMo 3世代”の設計思想 (長文・実務を意識した推論効率) が反映されている可能性は高い、という位置づけが妥当です (推定)。 ¹³

事前学習データと学習手法（公開範囲）

PLaMo 3.0 Prime β版の事前学習データについて、PFNは「独自データセット」に加え、**医療分野特化データセット**、NICTが整備する**日本語関連データセット**を学習に活用したと述べています。² ただし、総トークン数、データソース内訳（Web/書籍/論文/コード等）、言語比率（日本語・英語・他言語の割合）は**未公開／未特定**です。²

参考として、同じPFNブログで公開された「PLaMo 3 (NICT) 8B/31B」事前学習検証では、データ比率（英語43.75%・日本語31.25%・コード15%・多言語10%）と、学習トークン数（8Bで1T、31Bで3T）を明示しています。¹² ただしこれは“PLaMo 3系小規模事前学習モデル”の情報であり、PLaMo 3.0 Prime β版にそのまま適用できるとは限りません（Primeの比率・規模は**未公開**）。¹⁴

長文コンテキスト処理の工夫

PFNは、長いコンテキスト学習を事前学習段階から行うのは計算資源的に困難であるという一般的課題を踏まえ、事前学習後の「継続事前学習」で文脈長を拡張する方針を説明しています。¹⁰ 具体手法として、RoPE周波数スケージングの代表例である**YaRN**を採用し、コンテキスト長を**64K (=65,536) トークン**、出力側を**最大20Kトークン**まで伸ばしたと述べています（PLaMo 2.2 Primeは入力32K／出力最大4K）。¹⁰

“長考”を可能にする学習設計と推論制御

PFNは推論モデル（Reasoning Model）を「小さなステップに分割して段階的に解く」「探索的な試行錯誤や推敲の繰り返し」を行うよう訓練されたLLMと定義し、PLaMo 3.0 Prime β版で推論能力を中心に性能が向上したと位置づけます。⁵

学習パイプラインとしては、**SFT（教師あり）→DPO（直接選好最適化）→RL（強化学習）**を実施したと明記し、さらにSFT/DPOでは「最終回答だけでなく思考過程についても損失計算」する変更を加えたと述べています。¹⁰ RLについても、学習安定化の工夫を組み込み、ドメインごとに報酬関数を実装したことが記載されています。⁵

推論時の“長考”制御について、PFNは評価条件として **max_reasoning_tokens=12,288**（比較対象のQwen3-235B-A22B-Thinking-2507およびgpt-oss-120bは32,768）を設定したと開示しています。¹⁰ これは、少なくともPFNの推論基盤において「思考用トークンの予算」を明示的に制御できる設計であることを示唆します（ただしパラメータ名称・外部公開API仕様は**未公開**）。⁵

“長考”の出力表現として、PLaMo 3.0 Prime β版は「思考過程も含めて出力を生成する能力」を得たと示され、例題では思考→手順→根拠説明が整然と出力されています。¹⁰ ただし現状では、思考は主に英語で行われる（学習データ影響）とされ、正式版に向けて「思考を日本語で行うよう再訓練」を検討中で、トークン節約や表現ブレ低減を狙う、と明記されています。¹⁰

ツール連携（エージェント能力）の位置づけ

PFNプレスはReasoning機能により「ツール呼び出し能力」が向上したと述べ、PFN技術ブログでも「ツールの利用を適切に行えるようにするためのデータ」を追加したと説明しています。² 一方で、後述のBFCL（英語ツール利用）では最難条件（複数ターン×複数ツール選択）で競合に遅れがあることも同じく明示されており、“ツール連携で長考を拡張する”領域は改善余地が残る、が現時点の読みです。⁷

以下は、公表された「思考トークン」「ツール呼び出し能力」を前提に、企業実装で一般に採り得る“長考”処理フロー（想定）です。PLaMo 3.0 Prime自体がこの全工程を内包する、という意味ではありません（外部ツール実行・監査は実装側責務）。¹⁵

flowchart TD

A[ユーザー入力] --> B[プロンプト整形・ポリシー適用
PII/機密のマスキング等]
 B --> C{推論予算の設定
例: max_reasoning_tokens}
 C --> |短い/定型| D[非長考で回答生成]
 C --> |複雑/高リスク| E[長考: 思考過程の生成
分解→検討→仮説→検証]
 E --> F{追加情報が必要?}
 F --> |いいえ| G[最終回答の生成]
 F --> |はい| H[外部ツール/検索/RAGの呼び出し
※実装側で制御]
 H --> E
 D --> I[ログ・監査・評価]
 G --> I
 I --> J[ユーザーへ返却]

比較表

技術仕様の比較（公開情報ベース）

観点	PLaMo 3.0 Prime β	PLaMo 2.2 Prime	Qwen3-235B-A22B	gpt-oss-120b
開発形態	フルスクラッチ (PFN)	フルスクラッチ (PFN)	オープンウェイト (Alibaba Cloud ¹⁶ 系)	オープンウェイト (OpenAI ¹⁷)
アーキテクチャ詳細 (Dense/MoE等)	未公開 (「一新」とのみ)	未公開 (本件資料範囲)	MoE (総235B/活性22B、Experts 128/Activated 8)	Transformer + MoE (総117B/活性5.1B、Experts 128/Active Experts 4)
総パラメータ数	未公開	未公開 (本件資料範囲)	235B	117B
事前学習データ規模	未公開	未公開 (本件資料範囲)	約36T tokens (Qwen3全体の説明)	主に英語のテキスト (規模詳細は限定的)
言語比率 (日本語/英語等)	未公開	未公開	多言語 (100+言語)	「mostly English」
コンテキスト長 (入力)	65,536 tokens (64K)	32,768 tokens (評価設定)	32,768 native / 131,072 (YaRN)	128k

観点	PLaMo 3.0 Prime β	PLaMo 2.2 Prime	Qwen3-235B-A22B	gpt-oss-120b
出力上限 (本文明示)	最大20K	最大4K (PLaMo 2.2 Prime)	実装例では max_new_tokens 32,768	仕様上は可変 (reasoning effort等)
長考の制御	評価では max_reasoning_tokens=12,288	(不明)	thinking mode (enable_thinking)	reasoning effort (low/medium/high)
ツール連携	「ツール呼び出し能力」向上を強調	(不明)	Agentic (外部ツール統合を強調)	Web/Python/関数呼び出し等を設計に含む
ライセンス	未公開 (βモニター条件も未公開)	商用APIとして提供実績あり	Apache 2.0 (Qwen3全体の説明)	Apache 2.0
提供形態	βはモニター募集、商用は2026年6月中旬予定	商用 PLaMo Primeとして展開	配布/デプロイ (OSS基盤多数)	配布/デプロイ (OSS基盤多数)

出典：PFNニュースリリース／PFN技術ブログ／Qwen3公式 (Hugging Face・ブログ) ／OpenAI公式 (ブログ・モデルカード)。¹⁸

ベンチマーク比較 (PFN公開図から転記)

指標 (PFN図の11ベンチ)	PLaMo 3.0 Prime β	PLaMo 2.2 Prime	Qwen3-235B-A22B-Thinking-2507	gpt-oss-120b (reasoning effort=medium)
英語指示追従 (IFBench)	0.663	0.375	0.520	0.623
日本語指示追従 (JFBench)	0.371	0.279	0.301	0.323
英語ツール利用 (BFCL: 単一ツールのみ)	79.8	74.4	88.0	78.3
英語ツール利用 (BFCL: 複数ツールから選択)	77.2	71.6	81.9	73.3
英語ツール利用 (BFCL: 複数ターン×複数ツール選択)	41.0	4.0	66.0	60.5

指標 (PFN図の11ベンチ)	PLaMo 3.0 Prime β	PLaMo 2.2 Prime	Qwen3-235B-A22B-Thinking-2507	gpt-oss-120b (reasoning effort=medium)
QA長文 (LongBench v2)	0.348	0.298	0.382	0.427
QA医療 (MedRECT)	0.652	0.583	0.704	0.625
QA医療 (医師国家試験)	0.896	0.756	0.977	0.930
QA法令 (lawqa_jp)	0.943	0.857	0.971	0.943
英語対話 (MTBench)	8.97	8.05	8.71	8.88
日本語対話 (Japanese MTBench)	9.01	7.88	8.99	8.95
数学推論 (AIME 2024)	0.579	0.020	0.857	0.782
QA STEM (GPQA-Diamond)	0.521	0.307	0.790	0.703

注：表はPFN公開図（棒グラフ上の数値）をそのまま転記。評価条件の詳細（H100+vLLM、max_model_len、max_reasoning_tokens等）はPFNブログAppendix参照。⁷

推論速度・メモリ要件・推論コストの推定（根拠付き）

以下は「未公開 (=PLaMoのパラメータ数)」が残るため、**確度を明示した推定**です。推論速度は環境（GPU種類、並列度、量子化、KVキャッシュ設定、同時リクエスト数）で大きく変動するため、ここでは主に「活性パラメータ」「コンテキスト長」「思考トークン」の3点を代理指標とします。MoEでは推論計算量は活性パラメータに概ね依存するという前提は、近年の推論効率研究でも一般的に用いられています（例：MoEの活性パラメータと効率の関係を論じる分析）。¹⁹

観点	PLaMo 3.0 Prime β	Qwen3-235B-A22B	gpt-oss-120b
計算量の“目安”	未公開 (Dense/MoE/活性規模が不明)	活性22B params/token	活性5.1B params/token
重みメモリ (BF16 換算)	未公開	235B \times 2B \approx 470GB (計算)	117B \times 2B \approx 234GB (計算) だが、MoE重みをMXFP4量子化して単一80GBで動作を主張
実運用のGPU要件 (目安)	未公開 (PFN評価はH100で実施)	BF16のままなら複数GPUが現実的 (重みだけで470GB級)	「単一80GB GPUで効率的に動作」と公式に明記
“長考”の増分コスト	max_reasoning_tokens=12,288 (評価設定) \rightarrow 長考時にトークン増	thinking modeで < think > ブロック生成 \rightarrow 長考トークン増	reasoning effort (low/medium/high) でCoT長が増える (=計算増)

観点	PLaMo 3.0 Prime β	Qwen3-235B-A22B	gpt-oss-120b
速度上の推定	未公開 ：ただし思考が英語中心でトークン増になり得る（日本語化検討）	活性22Bのため、同条件ならgpt-ossより計算負荷が高い可能性（推定）	活性5.1B+量子化のため、同条件なら最も低レイテンシ寄りの設計（推定）

出典：Qwen公式モデルカードの活性パラメータ、OpenAI公式ブログ/リポジットリ/モデルカードの活性パラメータ・量子化・推論制御、PFNブログの評価条件（H100+vLLM、max_reasoning_tokens）および思考言語（英語中心）。²⁰

ユースケースと導入上の注意

PFNは、Reasoningモデルは従来生成AIより計算量が多く、応答に時間がかかる場合があるため、商用版の前にモニター企業を募集して「実際の利用環境でのトラフィック量や応答性能を検証」「フィードバックをもとに安定性や処理速度の改善」を行うと説明しています。² したがってモニター募集の狙いは、性能“精度”だけでなく、**推論コスト増を踏まえた運用妥当性（レイテンシ、同時処理、失敗率、品質ブレ）**を現場負荷で詰める点にあると解釈できます。²

想定業務領域は、PFNが向上点として列挙する「複雑な指示追従」「情報抽出・ロールプレイ等の対話」「数理・アルゴリズム」「ツール呼び出し」、およびPFN評価ベンチの構成（長文QA、医療QA、法令QA、RAG改善を意識したデータ等）から、少なくとも以下が中心です（※具体的な業務プロセス適用は企業側設計が必要）。²¹

- ・長文ドキュメントを前提とした業務：規程・契約・要件定義・議事録・調査資料などを同一プロンプトで扱う要約／論点抽出／整合性チェック。²²
- ・高信頼を要求する専門QA：医療・法律の質問応答（ただし“専門家の代替”ではなく、業務補助・一次ドラフト・検索補助としての位置づけが必須）。²³
- ・ルールと条件が絡む意思決定支援：複数条件整理→段階的推論→結論提示、という“思考プロセス”が必要な業務（申請審査、社内規程適合、問い合わせ一次対応の根拠提示など）。²

導入上の主な利点は、(1) 日本語指示追従・日本語対話で強い数値を示している点、(2) PFNが政府ガイドラインに沿ったAIガバナンス体制を掲げている点、(3) 既存PLaMo PrimeがクラウドAPIやオンプレ等多様な提供形態を展開してきた点（=将来の商用版でも同様の展開が期待されるが、PLaMo 3.0 Prime商用版の提供形態そのものは現時点で断定不可）です。²⁴

制約・注意点としては、(a) 長考は計算量増により遅延が増え得る、(b) BFCL最難条件・AIME・GPQAで課題が残る、(c) 思考が英語中心でトークン効率や表現ブレの課題があり得る、(d) β モニター条件・データ扱いが未公開、の4点が重要です。²⁵

モニター募集から商用化までの“公表情報に基づく”タイムライン（※6月中旬は代表日で図示）を以下に示します。²⁶

gantt

```
title PLaMo 3.0 Prime:  $\beta$ モニターから商用化まで（公表情報ベース）
dateFormat YYYY-MM-DD
axisFormat %Y-%m
```

section 研究開発の前提

```
PLaMo 3 (NICT) 小規模事前学習の検証公開 : milestone, 2025-11-14, 1d
```

section β公開と検証

PLaMo 3.0 Prime β 公開・モニター募集開始 : milestone, 2026-03-19, 1d
実運用での負荷/応答性能検証 (目的として明示) : 2026-03-19, 2026-06-10
安定性・処理速度改善 (目的として明示) : 2026-03-19, 2026-06-10

section 商用提供

商用版 正式提供 (予定: 6月中旬) : milestone, 2026-06-15, 1d

リスクと対策

安全性 (幻覚・誤答・過信)

Reasoningモデルは“思考プロセス”を提示できても、それ自身が正しさを保証しません。PFN自身も、PLaMo 3系事前学習モデルについて「不正確・偏った内容」や「不適切と判断される応答」を生成し得るため、用途に合わせた安全性テストとチューニングが必要と明記しています。²⁷ 医療・法令のような高リスク領域は、ベンチマークが高くても、実運用では根拠資料提示・人手レビュー・責任分界 (最終判断は専門家/業務責任者) を必須要件にすべきです。²⁸

対策: ①RAGで根拠を必須化、②回答の不確実性表明、③ドメイン別のレッドチーミング、④出力監査ログとリスクベース運用 (後述ガイドライン参照)、⑤モデル更新時の回帰テスト。²⁹

プライバシー・データガバナンス (入力データと学習データ)

企業利用では、プロンプトに個人情報・機密情報が混入しやすい点が最大論点です。日本の個人情報保護法は個人情報の定義を置き、取り扱いの適正化を求めます。³⁰ 個人情報保護委員会³¹ は生成AIサービス利用に関する注意喚起を公表しており、入力データの扱い (利用目的や安全管理等) を含む実務上の留意が必要です。³²

対策: ①入力段階でのPII検知・マスキング、②ログ/学習への二次利用を含む契約・設定確認、③データ保持期間・越境移転・委託先管理の整理、④監査可能なデータフロー (トレーサビリティ) 確保。³³

著作権・知財 (学習と生成の両面)

生成AIでは「学習段階」と「生成・利用段階」で論点異なること、また文化審議会での整理が公表されていることが重要です。文化庁³⁴ は「AIと著作権」に関する考え方をまとめ、関連文書 (チェックリスト等) も提示しています。³⁵

対策: ①学習データの出所・権利処理の説明可能性 (可能な範囲で)、②生成物の類似性リスクを下げるフィルタ・ポリシー、③利用規約での禁止事項 (権利侵害の誘発等) と監視、④社内のクリエイティブ/法務フロー統合。³⁶

バイアス・公平性

PLaMo 3.0 Prime β版の学習データ詳細は未公開のため、具体的バイアス源 (言語比率・領域比率・合成比率等) を外部から検証しづらい構造があります。² 参考としてPLaMo 3 (NICT) 事前学習検証では英語比率が最大で、日本語データ不足を高品質英語データの翻訳で補う戦略が説明されていますが、翻訳由来の分布偏りや表現癖が業務上のバイアスになり得ます (推定)。¹⁴

対策：①業務KPIに直結するバイアス評価（属性・地域・業界など）、②“根拠提示”を伴う回答への誘導、③重要業務は人手承認（human-in-the-loop）。³⁷

ツール連携・プロンプトインジェクション

PFNはツール呼び出し能力を基盤能力として重視していますが、ツール連携が進むほど、**プロンプトインジェクション**や“誤ツール実行”のリスクは増大します。³⁸

対策：①ツール実行権限を最小化、②実行前のルールベース検証、③“思考過程”をユーザーに開示しない運用（安全上の原則）、④監査ログと異常検知。³⁹

今後の展望

PFNは、モニター運用で得たフィードバックを踏まえて安定性・処理速度を改善し、2026年6月中旬に商用版を正式提供する予定だとしています。² また技術ブログでは、今後の方向性として「より長いコンテキスト長」「より高度なReasoning能力」「実務と密接に関わる領域タスクの性能向上」を掲げています。¹⁰ “思考の日本語化”も検討事項として明確に言及されており、トークン節約と表現ブレ低減が狙いです。¹⁰

国内での位置付けとして、PFNは「フルスクラッチで国内開発されたReasoningモデルとしては…国内初（当社調べ）」と明示しており、ここが“国産フルスクラッチ×長考（推論モデル）”の差別化点になります。³ この差別化が事業価値に転化する条件は、(1) “日本語の実務”で優位な評価軸（指示追従、対話、長文処理、医療・法令など）を継続的に伸ばすこと、(2) コスト高になりがちな長考を、推論予算制御・推論基盤最適化・必要時のみ思考を回す設計で吸収すること、(3) ガバナンス（透明性・リスク対応）をプロダクトに実装すること、の3点です。⁴⁰

国際競争との技術課題として注視すべきは、PFN自身が課題として名指しする「BFCL最難条件（複数ターン×複数ツール）」「AIME」「GPQA-Diamond」領域の底上げです。⁷ ここは、Qwen3が“thinking/non-thinking”のハイブリッド制御やエージェント能力を前面に出し、OpenAI gpt-ossがreasoning effortとツール学習（ブラウジング・Python・関数呼び出し）をモデル設計に含めている潮流と同型であり、“長考”を単体モデルの内省だけで完結させず、ツール・環境フィードバックまで含めたエージェント化へ接続できるかが次の分水嶺になります。⁴¹

出典

一次情報（優先） - PFNニュースリリース「生成AI基盤モデルPLaMo 3.0 Primeβ版のモニター企業募集」。³

- PFN技術ブログ「PLaMo 3.0 Prime β版をリリースしました」。²²

- ITmedia AI+「初の“長考”できる国産フルスクラッチLLM『PLaMo 3.0 Prime』…」。⁴² ※本文の完全取得は本環境制約があるため、確認できた抜粋範囲を一次情報で突合。²

補助情報（比較・ガバナンス・法務） - PFN技術ブログ「PLaMo 3 シリーズにおける8B, 31Bの小規模モデルによる事前学習の検証」。¹²

- Qwen公式（Hugging Face / 公式ブログ）。⁴³

- OpenAI公式（Introducing gpt-oss / gpt-ossリポジトリ / モデルカード）。⁴⁴

- PFN AIガバナンス。⁴⁵

- 経済産業省「AI事業者ガイドライン」。⁴⁶

- 個人情報保護委員会「生成AIサービス利用に関する注意喚起等」。³²

- 文化庁「AIと著作権について（考え方・関連資料）」。³⁵

重要引用URL (原文参照用)

<https://www.preferred.jp/ja/news/pr20260319>
<https://tech.preferred.jp/ja/blog/plamo-3-prime-beta-release/>
<https://www.itmedia.co.jp/aipius/articles/2603/23/news112.html>

https://tech.preferred.jp/ja/blog/plamo_3_8b_31b/
<https://huggingface.co/Qwen/Qwen3-235B-A22B>
<https://qwenlm.github.io/blog/qwen3/>
<https://openai.com/index/introducing-gpt-oss/>
<https://arxiv.org/pdf/2508.10925>
<https://www.preferred.jp/ja/company/aipolicy>

https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20241216_1.pdf
https://www.ppc.go.jp/news/careful_information/230602_AI_utilize_alert
<https://laws.e-gov.go.jp/Law/415AC0000000057>
<https://www.bunka.go.jp/seisaku/chosakuken/aiandcopyright.html>

1 12 14 27 https://tech.preferred.jp/ja/blog/plamo_3_8b_31b/

https://tech.preferred.jp/ja/blog/plamo_3_8b_31b/

2 3 11 15 18 21 25 26 31 38 40 <https://www.preferred.jp/ja/news/pr20260319>

<https://www.preferred.jp/ja/news/pr20260319>

4 5 10 13 16 17 22 34 <https://tech.preferred.jp/ja/blog/plamo-3-prime-beta-release/>

<https://tech.preferred.jp/ja/blog/plamo-3-prime-beta-release/>

6 7 23 24 28 <https://tech.preferred.jp/wp-content/uploads/2026/03/image1.png>

<https://tech.preferred.jp/wp-content/uploads/2026/03/image1.png>

8 9 42 <https://www.itmedia.co.jp/aipius/articles/2603/23/news112.html>

<https://www.itmedia.co.jp/aipius/articles/2603/23/news112.html>

19 <https://arxiv.org/abs/2603.17280>

<https://arxiv.org/abs/2603.17280>

20 41 43 <https://huggingface.co/Qwen/Qwen3-235B-A22B>

<https://huggingface.co/Qwen/Qwen3-235B-A22B>

29 33 37 46 https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20241216_1.pdf

https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20241216_1.pdf

30 <https://laws.e-gov.go.jp/law/415AC0000000057>

<https://laws.e-gov.go.jp/law/415AC0000000057>

32 https://www.ppc.go.jp/news/careful_information/230602_AI_utilize_alert

https://www.ppc.go.jp/news/careful_information/230602_AI_utilize_alert

35 <https://www.bunka.go.jp/seisaku/chosakuken/aiandcopyright.html>

<https://www.bunka.go.jp/seisaku/chosakuken/aiandcopyright.html>

36 https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/seisaku/r06_02/pdf/94089701_05.pdf

https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/seisaku/r06_02/pdf/94089701_05.pdf

39 <https://github.com/openai/gpt-oss>

<https://github.com/openai/gpt-oss>

44 <https://openai.com/index/introducing-gpt-oss/>

<https://openai.com/index/introducing-gpt-oss/>

45 <https://www.preferred.jp/company/aipolicy>

<https://www.preferred.jp/company/aipolicy>