# 知財業務におけるPLaMo 3.1 Prime 導入設計図

機密性・コスト・推論力の限界を見極める、AIアーキテクチャ戦略

対象モデル: PLaMo 3.1 Prime (2026年10月27日リリース)
評価ベース: 独自調査および公式技術情報

Gemini Notebook

# 結論：PLaMo 3.1 Primeの知財業務における位置づけ

## 最適な用途 - 閉域での大量処理

**核心:**
秘密情報を社外に出さずに、日本語文献を読む・まとめる・調べる用途において極めて有力な選択肢。

**主要機能:**
「オンプレミス環境」と「Deep Research」の組み合わせによる、社内文書限定の閉域調査。



## 不適な用途 - 高度な法的推論

**核心:**
クレーム解釈、拒絶理由への反論、FTO調査など、誤りの代償が大きく高度な法的推論が求められる業務には代替不可。

**限界:**
STEM・コーディング領域において、同クラスのオープンモデルを下回る（PFN自社測定）。

【基本スペック】提供開始: 2026年10月27日 / API標準料金: 入力60円・出力250円（100万トークンあたり）

Gemini Notebook

# 開発系譜とソブリンAIとしての信頼性



## GENIAC & NICT
### 官民連携の開発基盤

GENIAC第1サイクルでPLaMo-100Bを開発。NICTとの共同研究による事前学習・安全性評価。設計から学習まで国内完結。



## デジタル庁「源内」
### ガバメントAI採択

2.0 Primeが試用モデルに選定。PLaMo翻訳は2025年12月より稼働中。2025年日経優秀製品・サービス賞最優秀賞受賞。



## 日本総研 全社導入
### エンタープライズ実稼働

3.0 Primeをオンプレミスで全社利用（設計書レビュー、規程検索など）。約半年の評価を経て本番稼働。

⚠ 【調達時の確認事項】SOC2およびISMAPの取得状況については、第三者記事で未取得との記述があるため、導入要件に含まれる場合はPFNへの直接確認が必須。

Gemini Notebook

# コア・イノベーション：ハイブリッド推論とDeep Research

**事実性・忠実性の向上**

ハルシネーション関連8ベンチマークで3.0 Prime比 約15%改善。知らないことは「分からない」と答える事後学習の強化。



PLaMo 3.1 Core

**非推論モード**

単純タスク用。出力トークン数を約90%削減し、前世代比で最大1.7倍の速度向上（投機的デコーディング）。大量の公報要約に最適。

**推論モード / Deep Research**

調査計画の立案から出典付きレポート作成まで数分〜数十分で自律実行。

詳細レポート

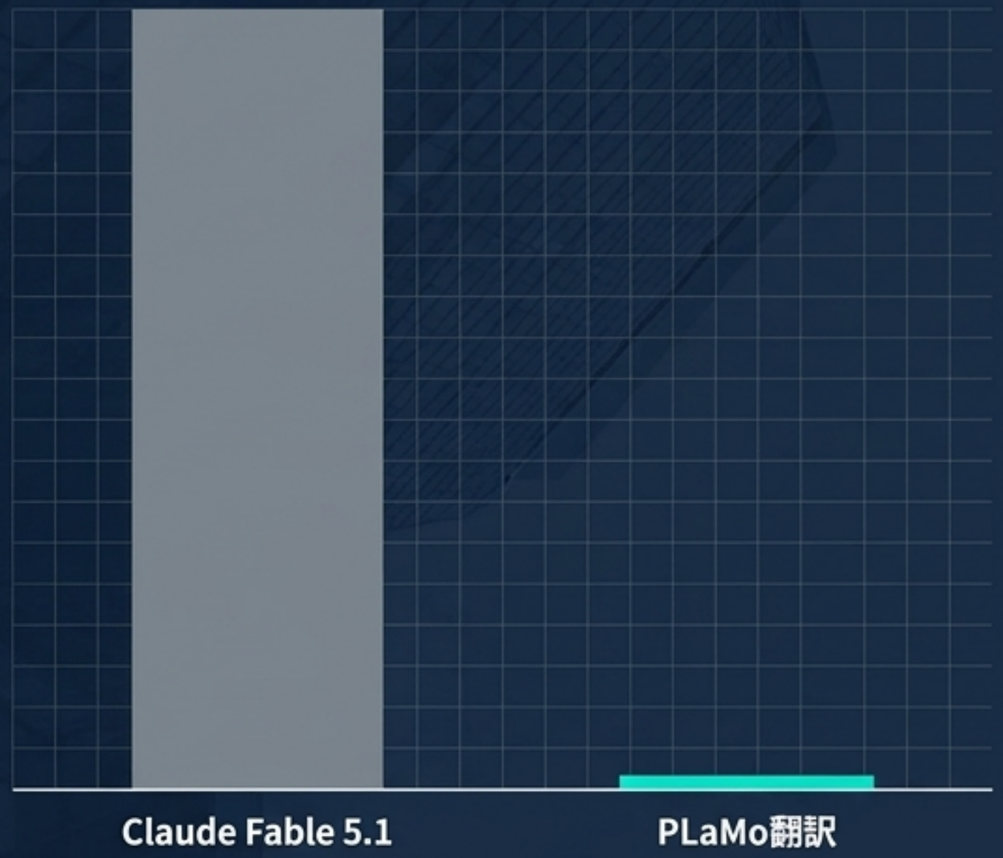
Gemini Notebook

# LLMランドスケープ比較マトリクス

| | PLaMo 3.1 Prime | 海外フロンティア（GPT/Claude/Gemini） | 他の国産LLM（tsuzumi, Sarashina3等） |
|---|---|---|---|
| 安全性・閉域運用 | 比較対象中で最高。オンプレミス対応。 | データの越境、経済安全保障上の懸念。 | 国内完結。 |
| 日本語特化 | 非常に優秀。 | 汎用的に優秀だが特化型ではない。 | tsuzumi（軽量30B・知財特化版あり）、Sarashina3（30兆トークン学習）。PLaMoのような知財特化のパッケージ事例は現状未確認。 |
| 高度な推論 | STEM・コーディング・Web検索で同クラスのオープンモデルに劣後。 | 総合的な推論力、長文処理、複雑なロジック構築に圧倒的強み。 | |
| コスト | 入力60円・出力250円（1Mトークン） | モデルにより高額。 | |

Gemini Notebook

# 隠れた最大の武器：PLaMo翻訳（plamo-3-translate）




## 圧倒的なコストパフォーマンス

性能で上回ったのはClaude Fable 5.1のみ。しかし、コストはその約85分の1（比較対象API全体の平均でも約30分の1のコストで稼働）。




## 特許文体への自律的適応

ベースモデル（最大31B）の学習データに占める日本語割合は約30%。英文クレーム入力時、文脈から特許文書と判断し、「前記」などの特許特有の文体で自然に出力する能力を確認（PFN技術ブログより）。

【ユースケース】
外国出願向け翻訳、外国公報の大量和訳におけるコスト破壊。

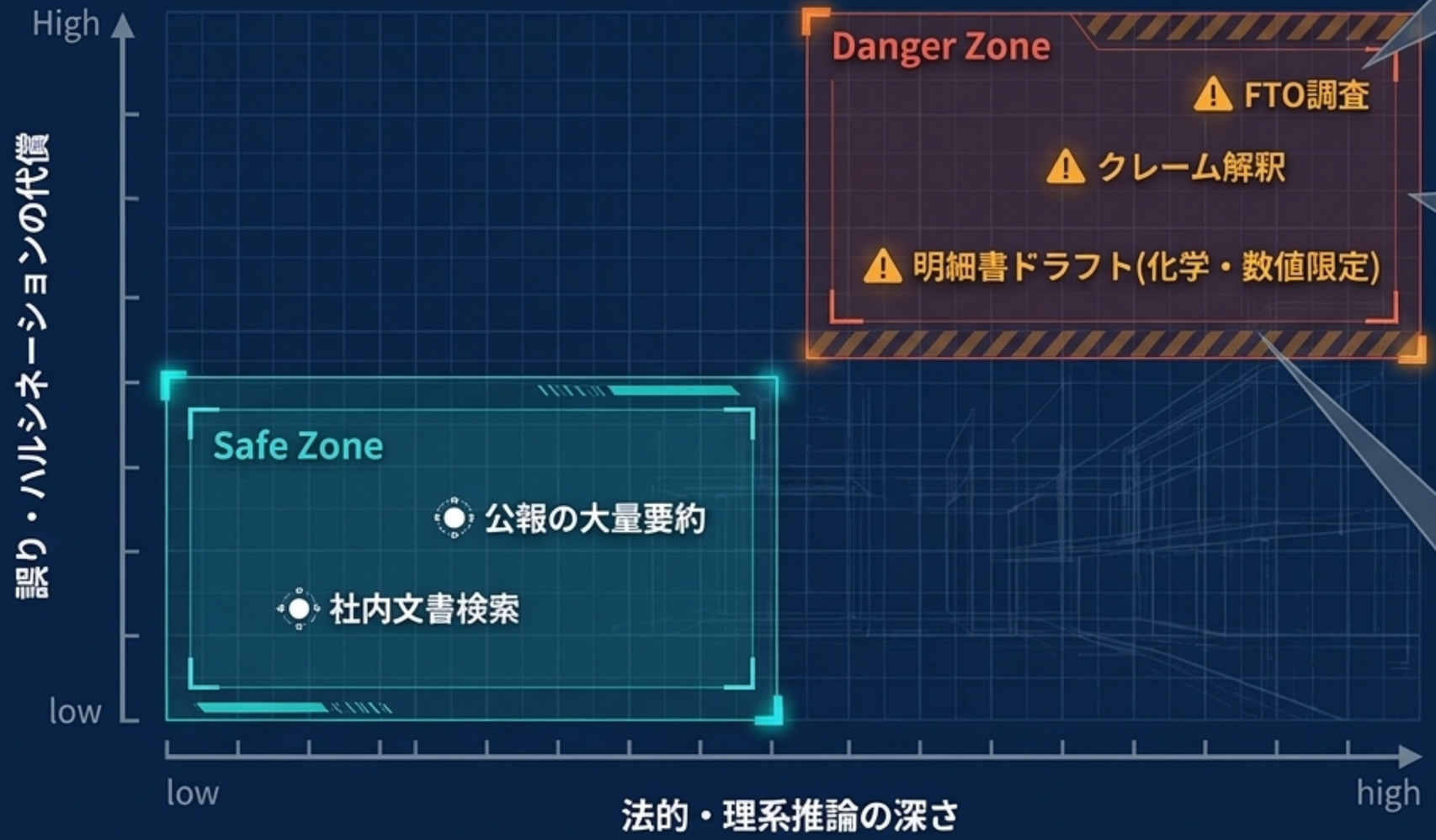
Gemini Notebook

# 知財業務適性ヒートマップ

| 知財タスク | 適性 | 人の依存度 | 理由・コンテキスト |
|---|---|---|---|
| 特許公報の読解・要約 | ◎ | 低 | 非推論モードで安価・高速。<br>一次スクリーニングに最適。 |
| 外国公報の和訳 | ◎ | 中 | PLaMo翻訳。用語統一やクレーム限定関係の欠落チェックは人が必須。 |
| 社内技術文書の横断調査 | ◎ | 低 | オンプレミス閉域網での最大の利点。 |
| IPランドスケープ（非特許情報） | ○ | 中 | 市場・技術動向の調査。特許統計は別途DBから与える必要あり。 |
| 明細書ドラフト作成 | △ | 高 | 実績なし。下書き用途に限定。 |
| 中間処理（進歩性反論） | △ | 高 | 引用文献の対比整理には役立つが、反論ロジックは作れない。 |
| クレーム解釈・FTO調査 | × | 極高 | 法令分野の弱さと推論力不足。誤りの代償が大きすぎる。 |

Gemini Notebook

# 限界のメカニズム：なぜFTOやクレーム解釈に用いてはならないのか




## STEMの弱点

化学構造式、数値限定、パラメータ発明の理解において致命的な誤読リスク（公式評価で競合を下回る）。

## 法令解釈の壁

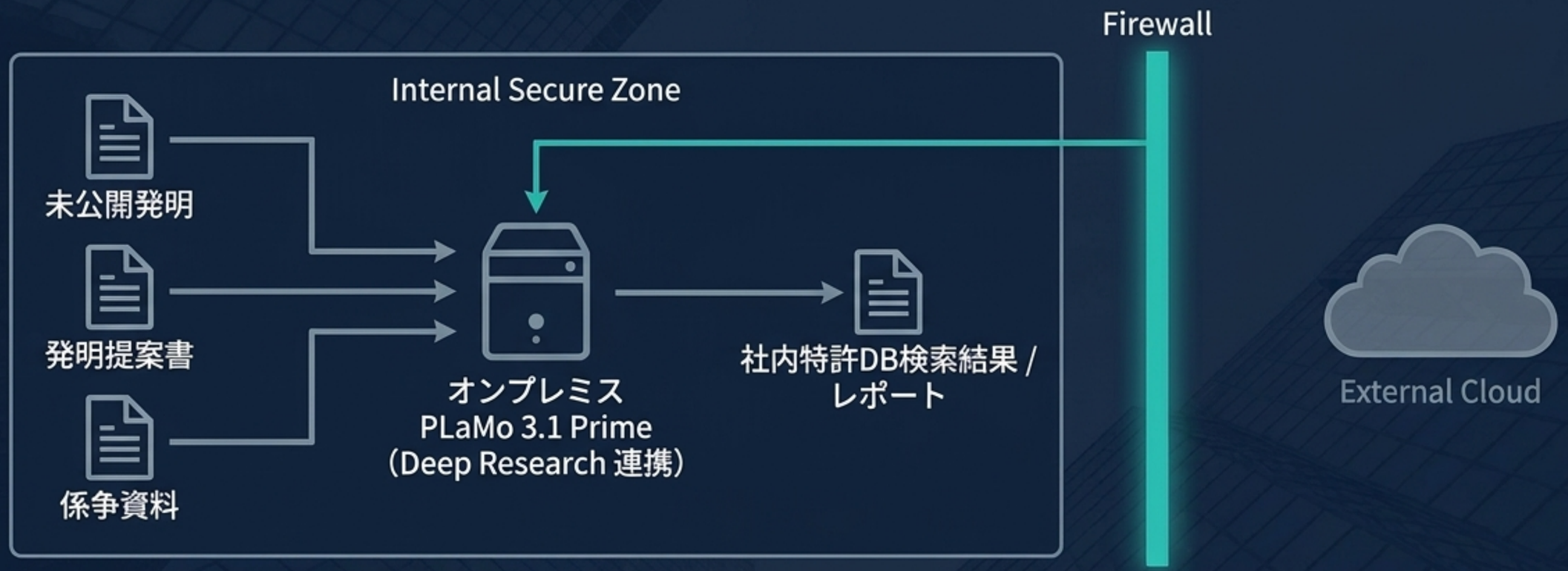
3.0時点で日本の法令分野が弱点。最新の判例や審査基準の適用において、一次資料の照合なしでの法的判断は不可。

## 網羅性の欠如

Deep Researchは指定コーパスには強いが、J-PlatPat等への特許DB直結が公式にJ-PlatPat直結が公式に確認されておらず、母集団形成には使えない。

Gemini Notebook

# 戦略的優位性：オンプレミス環境による新規性喪失リスクの完全遮断




## 営業秘密管理の徹底

未公開情報を外部APIに送信する心理的・実務的ハードルを排除。社内承認プロセスが大幅に短縮される。

## 新規性喪失の防止

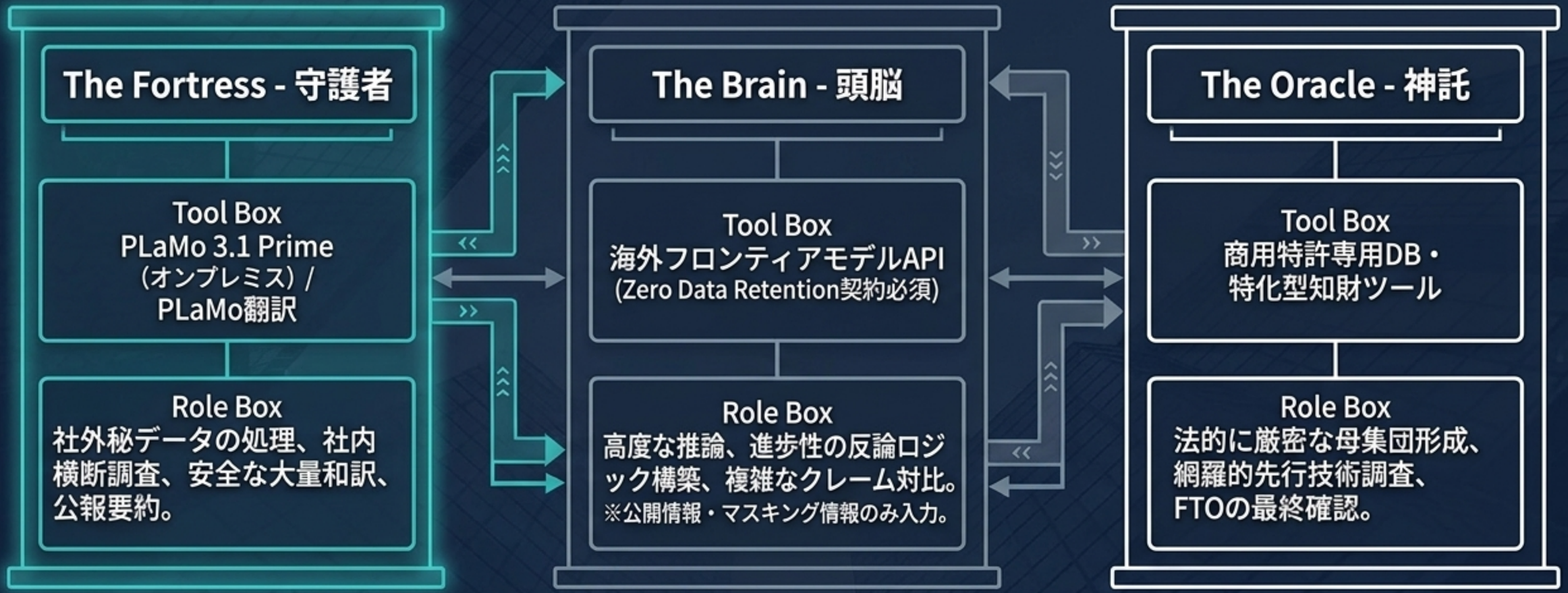
クラウドAPI入力が直ちに「公知」とはならないものの、Freeプラン等での学習利用リスクを根本から断つ。

## 導入の要件

GPUインフラ等の運用負荷がかかるため、部門単独ではなく全社AI基盤としての導入（日本総研モデル）が現実的。

Gemini Notebook

# 最適解：3系統によるIP・AIアーキテクチャの役割分担




**【結論】PLaMoは万能ツールではない。しかし、知財部門が長年抱えていた「機密の壁」を突破する、アーキテクチャの最重要ピースである。**

Gemini Notebook

# 導入に向けた実装プレイブック（Next Steps）

## 01 PoCの設計と品質検証

過去の自社案件（公開済みの公報や応答済みの拒絶理由）を使い、要約の欠落率と出典の正誤率を測定する。

Freeプランは入力データが学習に利用される公式説明があるため、機密情報のテスト入力は厳禁。

## 02 システム要件とコストの確認

PFNに対し、Deep Researchにおける自社特許DB（OpenSearch等）との接続可否、およびWeb検索時の追加課金体系を書面で確認する。

## 03 社内ガイドラインの改定

「オンプレAI」「外部API」「特許DB」の3系統の使い分け原則を策定。FTOや法的判断への生成AI利用禁止を明文化する。

Gemini Notebook

# Key Takeaways

## Blueprint Summary Card



**核心価値:** PLaMo 3.1 Primeの真価は、オンプレミスでの社内文書特化型Deep Researchと、高精度な特許翻訳（圧倒的低コスト）にある。



**限界の認識:** 法令解釈や複雑な理系推論（FTO、クレーム解釈）は現状の能力境界外であり、海外フロンティアモデルや専用DBとの併用が必須。



**戦略的配置:** 万能な単一ツールを求めるのではなく、「機密性を担保する基盤（Fortress）」としてIPシステム全体に組み込むことが導入成功の鍵となる。

Gemini Notebook