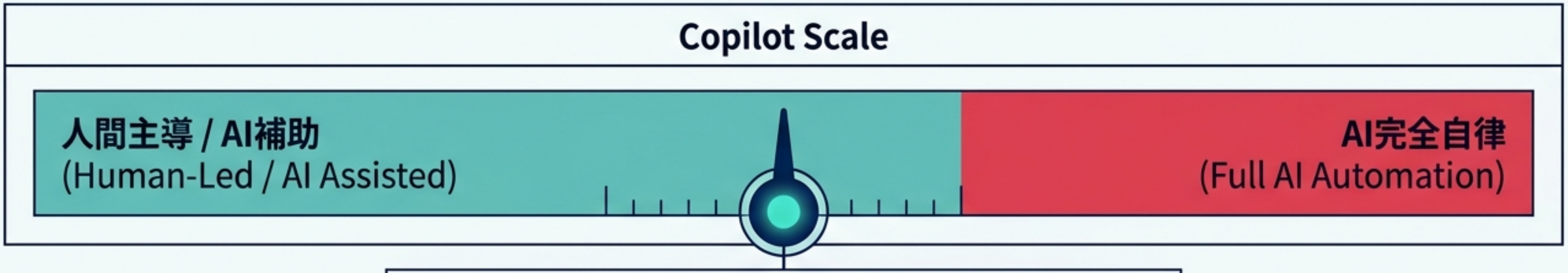
知財部門長・企業法務・弁理士・知財IT戦略担当者向け評価レポート

ソフトバンク「Sarashina」の知財業務適性評価と導入ロードマップ

「知財オートパイロット」から「知財コパイロット」への軌道修正と、安全なシステム統合アーキテクチャ

結論：厳格な人手審査を前提とした「知財コパイロット」としての導入を推奨



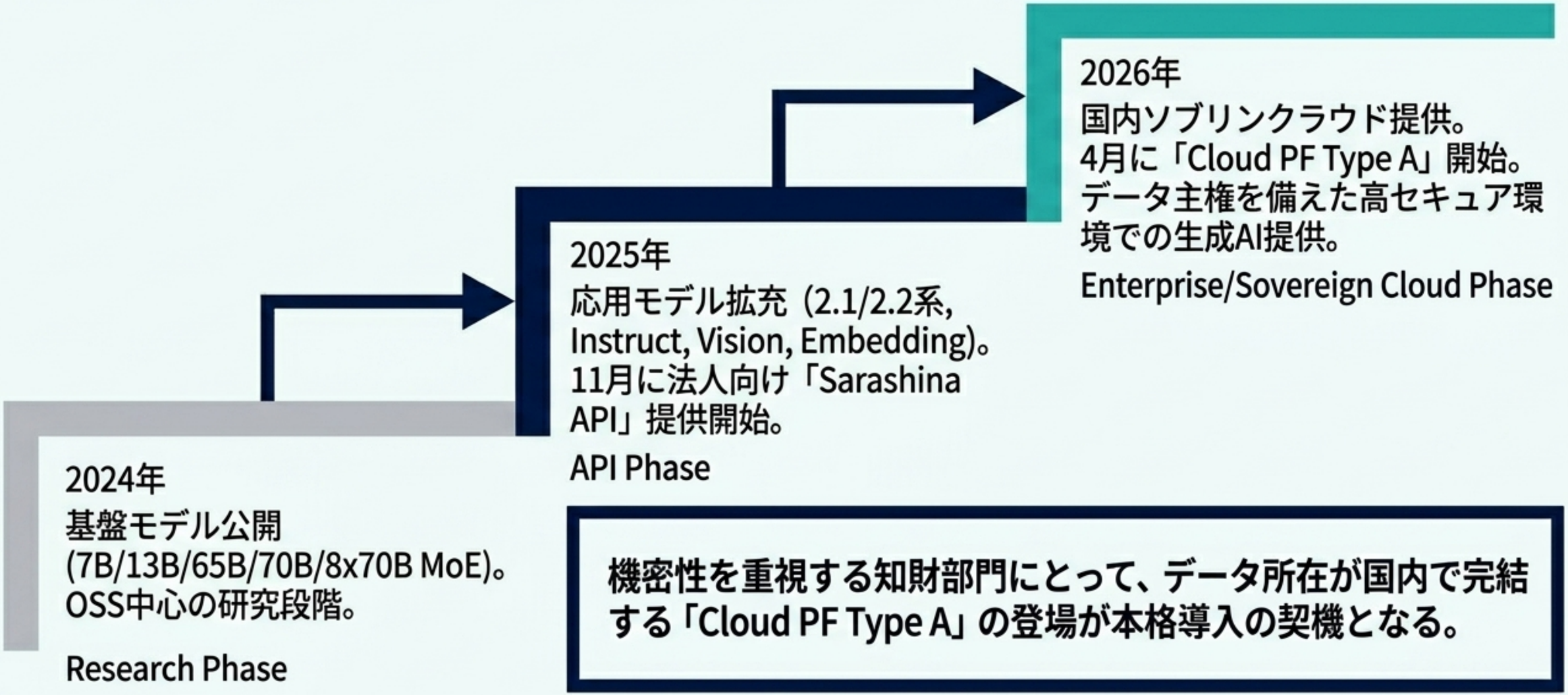
推奨 (Recommended)

- **用途:** 要約・翻訳、文書構造化、RAG検索補助、出願ドラフト補助
- **技術方針:** Sarashina API または MITライセンス系OSSの利用
- **環境:** データ主権を担保するCloud PF Type Aの活用

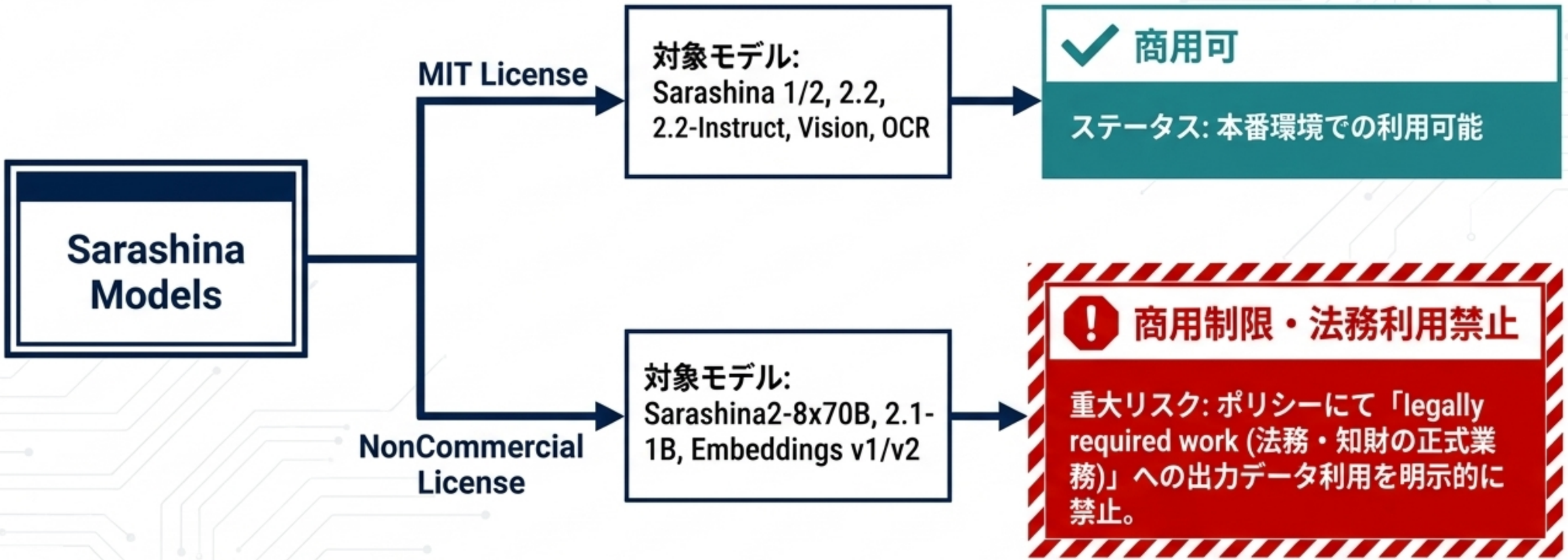
非推奨 (Not Recommended)

- **用途:** 侵害リスク評価、登録可否判断、法的見解の単独提示
- **技術方針:** 非商用(NonCommercial)ライセンス系OSSの本番利用
- **リスク:** 「legally required work」での利用規約違反、ハルシネーションによる法的誤誘導

研究開発基盤から「エンタープライズ・ソブリンクラウド」への進化



ライセンスの地雷原：OSSモデルの「商用・法務利用」に関する厳格な境界線



エンタープライズの回避策 (Enterprise Alternative):
Sarashina API / Cloud PF Type A は商用サービスであり、この制限を回避する「安全なバイパス」として機能する。

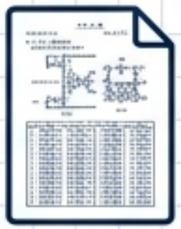
モデル別知財業務適合性マトリクス

系列	主な仕様	ライセンス	知財実務評価
Sarashina2.2-Instruct (3B)	軽量・対話用。知財コパイロットのPoC第一候補。	MIT	✓ 推奨
Sarashina2.2-OCR & Vision	日本語縦書き・複雑な知財PDFの前処理に極めて有用。	MIT	✓ 推奨
Sarashina2-8x70B	研究用。法務本番利用不可。インフラ要件過大。	NonCommercial	✗ 不可
Embeddings v1/v2 (OSS)	高性能だが商用制限あり。そのままの知財RAG利用は不可。	NonCommercial	✗ 不可
Sarashina API (商用版)	知財部門の第一選択肢。入力データの学習不使用を明記。	商用利用可	✓ 最優先

物理からデジタルへ：知財ワークフローを牽引するVisionとOCR性能

入力課題

古い公報の
縦書きPDF



複雑な表を含む
特許明細書

図面番号を含む
証拠画像

Sarashina2.2-OCR
（CER 22.6達成）
/
Sarashina2.2-
Vision-3B

出力価値

高精度のMarkdown/
HTML/LaTeX再構成。

表崩れや図面番号
の取り違えを防止。

構造化データによる
高度な検索・分析

**「紙から構造化データへ」の変換において、他社LLMを凌駕する
日本語特有のレイアウト解析力を発揮。**

Digital Dossier

業務シナリオ別診断：AIに任せるべきタスクと、人間が守るべき領域

強く推奨 (Safe)

要約・翻訳・PDF構造化

- 推奨構成: API, 2.2-Instruct, OCR
- 主なリスク: 固有名詞の誤訳

補助用途で可 (Caution)

契約レビュー

- 推奨構成: APIでの条項抽出・差分比較
- 主なリスク: 業法・裁判実務の見落とし

条件付き可 (Caution)

先行技術調査

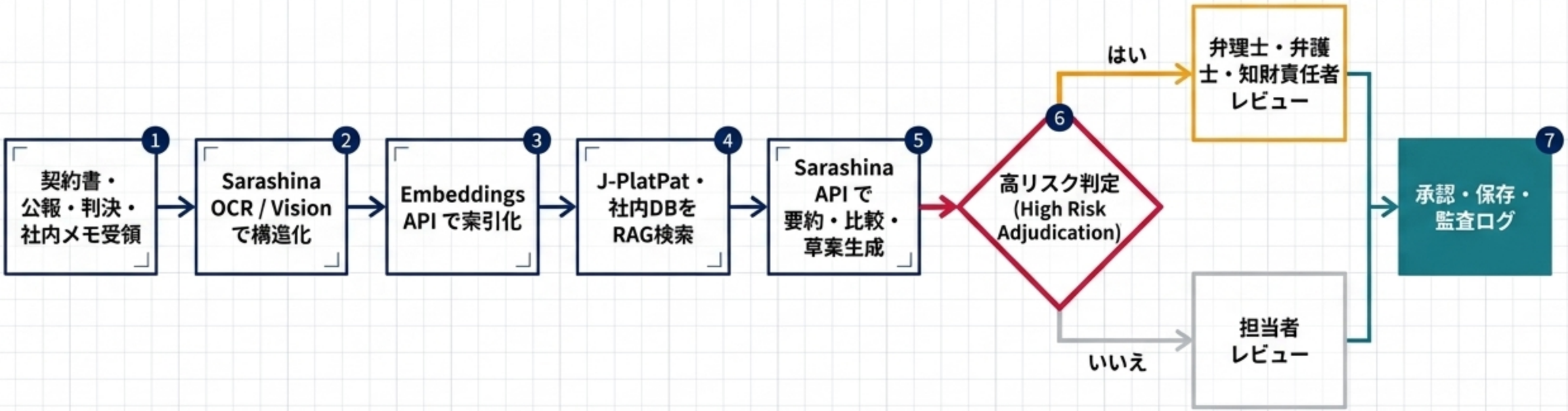
- 推奨構成: J-PlatPat主導、AIはクエリ展開・要約のみ
- 主なリスク: 検索漏れ、exhaustive recallの欠如

単独利用は不可 (Danger)

侵害リスク評価・最終クレーム確定

- 推奨構成: 人間主導限定
- 主なリスク: 誤ったクレーム解釈による重大な法的誘導

セキュア知財RAGアーキテクチャ：「人間主導審査」を組み込んだシステム統合



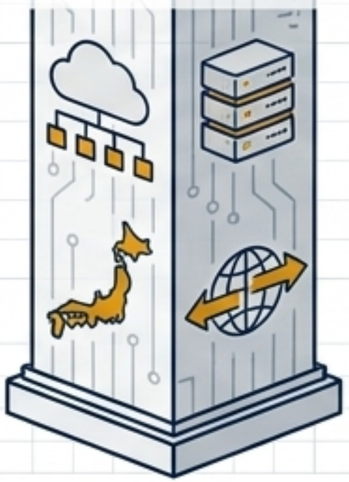
運用上のブラインドスポット：PoC移行前に解消すべき「未確認事項」

著作権と学習データ証明

日本語コーパスの具体的サイト、著作権者対応手順、文化庁ガイドラインに基づく確定的評価の不在。

インフラとデータ主権

Cloud PF Type Aは国内完結だが、標準APIの再委託先や越境移転の有無は要確認。



運用SLAとログ統制

ログ保持期間、削除証跡、可用性保証(SLA)、障害時補償の公開情報不足。

Web参照機能の制御

SearchBotによるWeb参照機能の「完全オフライン化」および無効化制御の可否。

知財特化型PoC（概念実証）評価ダッシュボードとデータ戦略

精度（Accuracy）

汎用ベンチマークではなく「知財ゴールデンセット」での再現性を評価。

指標：Recall@20、Hit@5、引用根拠付与率、ハルシネーション(幻覚)率。

効率（Efficiency）

レビュー工数削減率（目標20〜40%削減）

重大な誤読・事実誤認ゼロ。

セキュリティ (Security)

機密情報(PII/営業秘密)の誤送信0件

アクセス権逸脱0件。

段階的データ投入戦略

初期検証は「公開公報・雛形契約」に限定し、
機微情報（発明者名、未公開発明）はマスキングを徹底。

契約締結の絶対条件：ソフトバンクへ要求すべきエンタープライズ要件

顧客データの二次利用禁止の明文化

入出力、埋め込み、ログをAI再学習に一切使用しない旨の契約文書化。

保持・削除条項と監査権

ログ・キャッシュの保持期間の明記、削除証跡の提供、および監査質問権の確保。

モデル変更管理とSLA

基盤モデル更新時の事前通知、性能影響の説明、応答時間目標(SLO)の合意。

閉域接続・分離環境の確約

高機密案件向けのVPC/専用線接続、またはCloud PF Type Aにおける専有環境の提供。

最終戦略方針：知財コパイロットとしての安全な実装ルート

【本命ルート】Sarashina API / Cloud PF Type A

✓ データ主権と商用保証を確保したエンタープライズ統合の第一選択肢。
契約によるSLA/ログ統制の補完が必須。

【補完ルート】MITライセンス系OSS (2.2-Instruct, Vision, OCR)

社内閉域環境での文書前処理・構造化に限定して活用。

【利用不可】NonCommercial系OSS (8x70B, Embeddings)

✕ 法務・知財の正式業務(legally required work)への適用はコンプライアンス上、厳格に除外する。

「AIの能力」だけでなく、「情報の出所統制・削除統制・監査可能性」を設計の主軸に据えること。