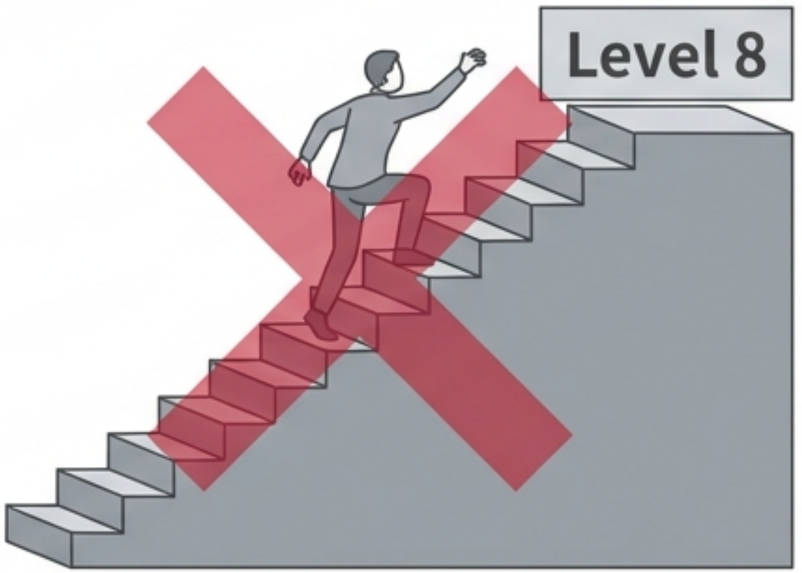
知的財産部門における AI導入の8段階

知財実務への翻訳と、日本企業のための「レベル・ポートフォリオ」戦略

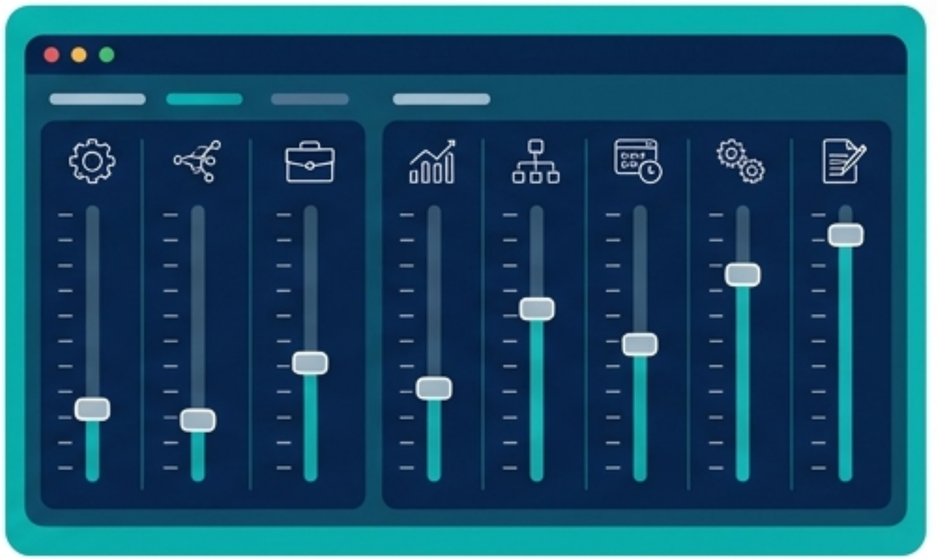
独自の実務的整理｜対象：知的財産部門 責任者・知財戦略担当者

目指すべきは、部門全体の「レベル8」化ではない



誤ったアプローチ

「AIモデルの性能競争」や「組織の成熟度」と捉え、全業務を一律に最高レベルのAIへ移行させようとする。



正しいアプローチ

業務ごとの「AIへの委任量」と「誤りの影響」を照合し、最も合理的なレベルを組み合わせる。

結論：知財部門を一つの数字で評価せず、業務別の「レベル・ポートフォリオ」として再設計することが真の知財DXである。

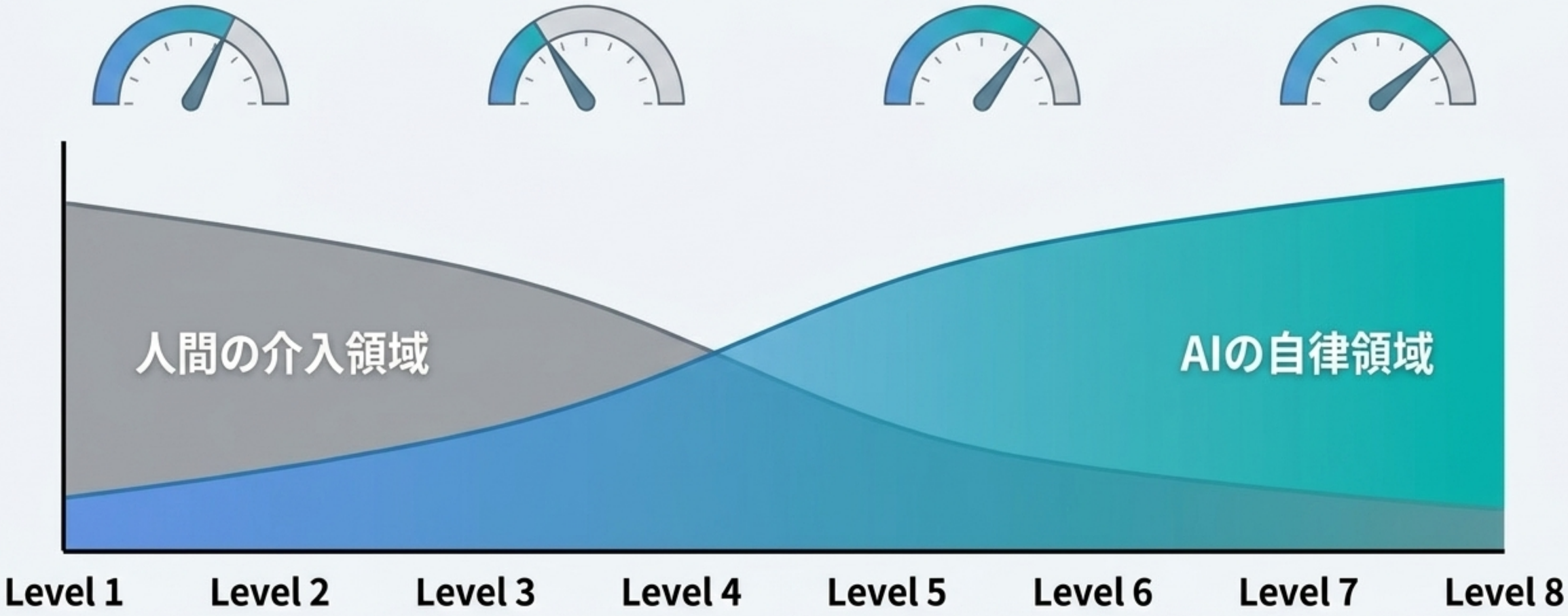
原典「8段階モデル」が測定する4つの次元

AIへの委任範囲

途中の人間介入

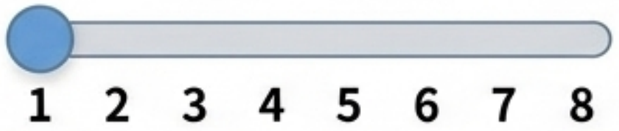
自律的開始の度合い

エージェント管理主体



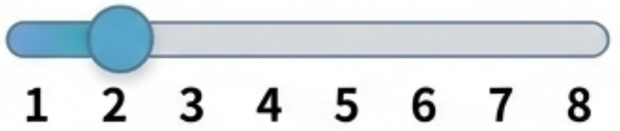
この8段階は「優劣」ではなく「委任のアーキテクチャ」を示す。法定期限と不可逆性を伴う知財実務では、高レベル化しつつも「重要判断には承認ゲートを残す二層構造」が適する。

人間主導の自動化：日常的な知財実務のスイートスポット (Level 1–3)



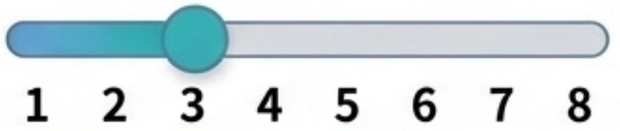
Level 1: Chatbot (問答型)

- 定義: 人が毎回課題を与え、AIが回答を返す。作業仮説の提示。
- 知財の適用例: 公報・判決の要約、検索観点のアイデア出し、クレームの平易化。
- 人間の役割: 質問設定、事実・法的評価の全面確認。



Level 2: Copilot (伴走型)

- 定義: AIが作業環境に入り、人と同一成果物を共同編集する。
- 知財の適用例: 明細書・意見書の推敲支援、特許一覧の共同作成、検索式の対話的拡張。
- 人間の役割: 作業の主導権維持、その場での採否判断。



Level 3: Agent (承認付き実行型)

- 定義: AIが複数工程を順に進め、重要な分岐で承認を求める。
- 知財の適用例: 発明届から調査報告書の連続作成、拒絶理由通知の論点抽出から反論案作成。
- 人間の役割: 承認ポイントの設定、専門的判断（クレーム解釈等）の実施。

システム主導の実行：定型・監視業務の自律化（Level 4–6）

メカニズム: 限定業務を完遂させ、最後に結果を確認。

知財タスク: 競合出願の一次分類、書誌・ファミリー情報の整備。

メカニズム: 指示を待たずバックグラウンドで監視・通知。

知財タスク: 競合・技術・判例の常時監視、期限接近・異常変化の自動通知。

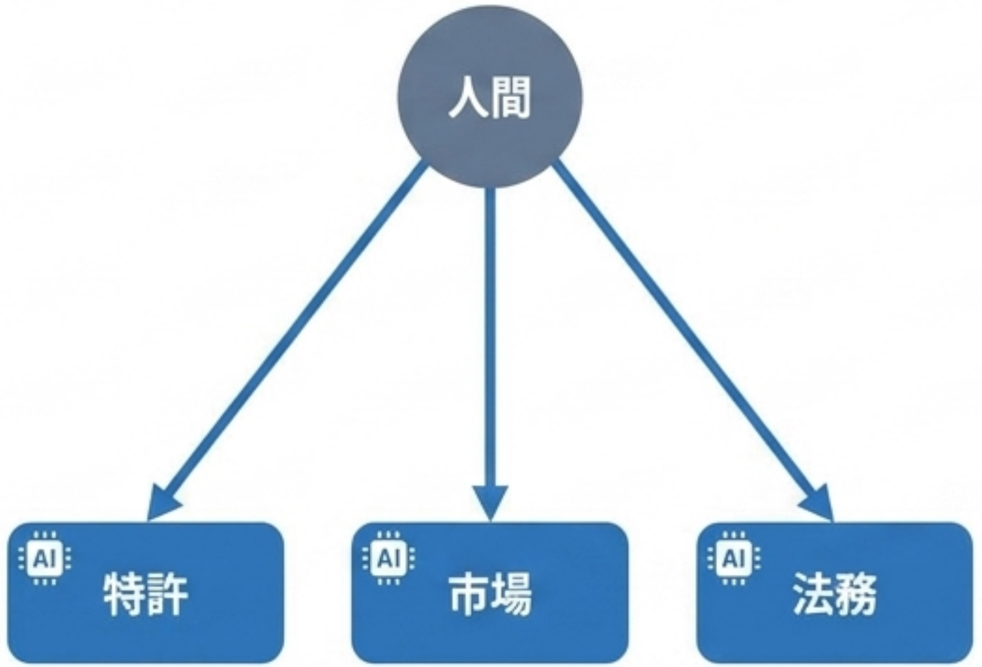
メカニズム: RAG、社内基準、監査ログを統合し、再現可能な仕組みへ。

知財タスク: 発明評価・出願判断資料作成の標準ワークフロー化。

欠陥のある業務プロセスをそのまま自動化すると、誤りも大規模化する。
先に業務自体を整流化することが高レベル化の前提。

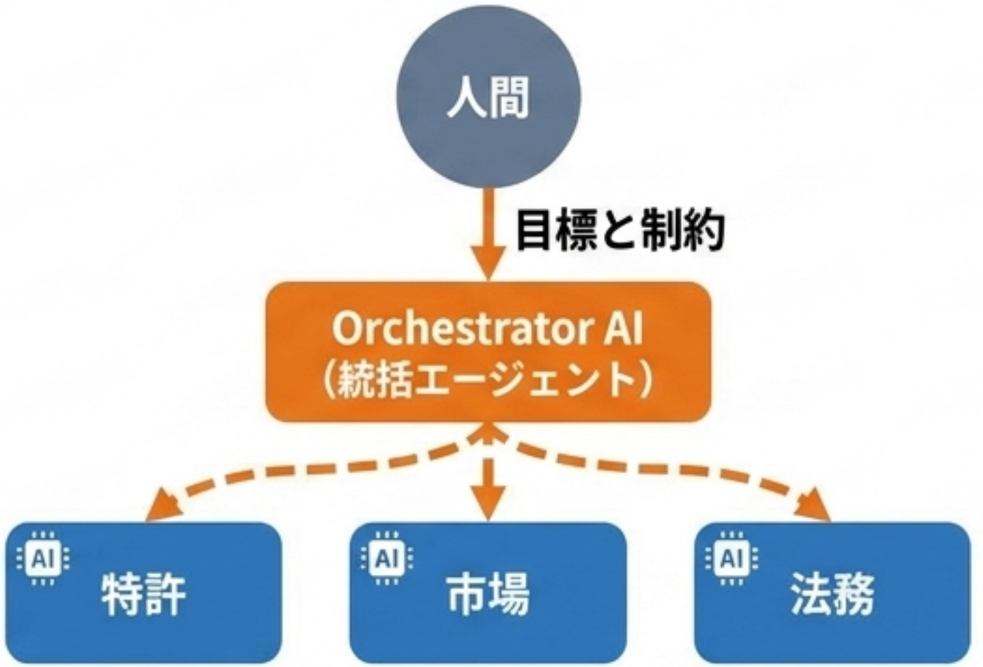
オーケストレーションの境界：Level 7 と Level 8 の決定的違い

Level 7: Multi-agent（複数エージェント型）



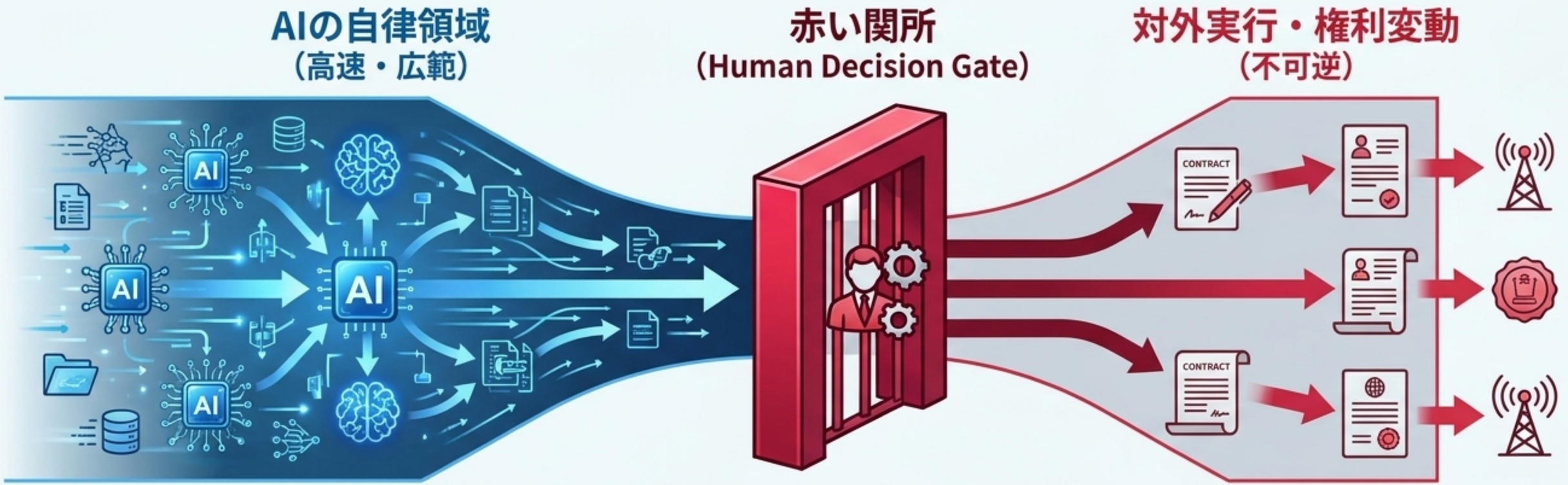
- **管理主体**：人間が直接管理
- **知財での意味**：人間が「特許」「市場」「法務」のエージェントに個別に指示を出し、人間が結果を統合・相互検証する。

Level 8: Orchestrator（統括エージェント型）



- **管理主体**：管理エージェント（AI）
- **知財での意味**：人間は「目標と制約」のみを与え、管理AIが課題分解・配分・矛盾解消・結果統合を一任される。人間の役割は「システム設計と最終審査」へ移行。

中核原則：「分析の自律度」と「意思決定権限」の完全分離



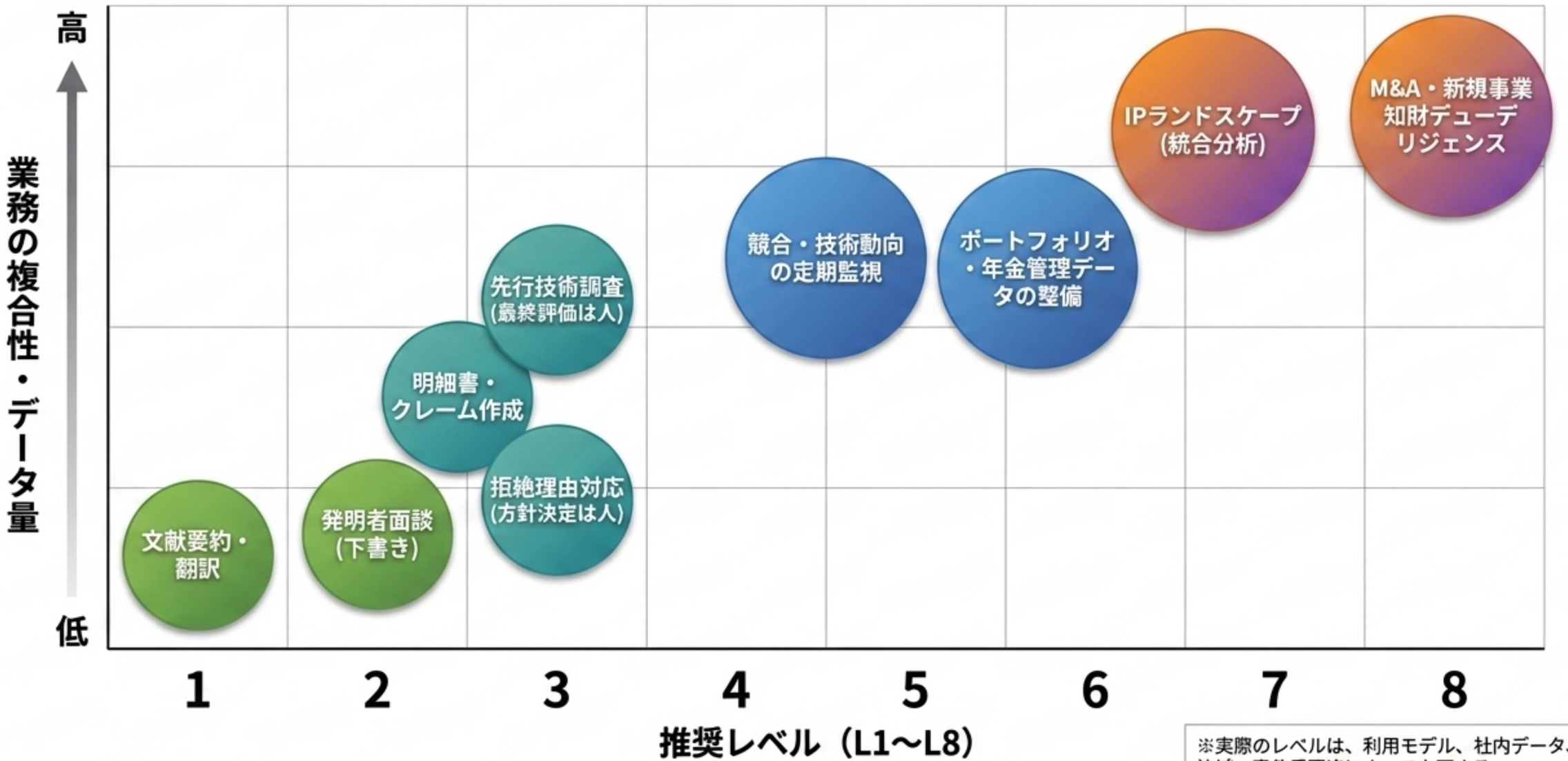
広く調べ、比較し、選択肢を生成する（先行技術調査、IPランドスケープ、反論候補作成）。
ここではLevel 5〜8の高度な自律性を許容。

人間による明示的な承認・専門的評価。

権利・法的地位・対外関係を変えるアクション。
AIには決して一任しない。

高リスク業務を低機能なAIだけで行うという意味ではない。高度なAIで分析させつつ、意思決定のインターフェースだけを承認付き（Level 3相当）に戻すアーキテクチャが必須である。

知財業務ポートフォリオ：適正レベル・マッピング



※実際のレベルは、利用モデル、社内データ、法域、案件重要度によって上下する。

The Red Zone：AIに一任してはいけない不可逆的業務

以下の業務は、誤りの影響が甚大であり、不可逆性または対外的拘束力を持つため、人間の明示的承認（意思決定ゲート）を必須とする。

権利の発生と消滅:

- 出願するか、ノウハウ・営業秘密として秘匿するかの決定
- 特許・商標等の維持、放棄、譲渡、担保設定

権利範囲と帰属の確定:

- 発明者、共同出願人、権利帰属の確定
- クレーム、補正、分割出願、外国出願の最終決定

法的見解と係争・契約:

- 非侵害、無効、FTOに関する最終的な法的見解
- 警告状、異議、無効審判、訴訟、和解の開始・終了
- ライセンス条件、共同研究・秘密保持条件の決定と契約締結
- 経営資源配分を伴う知財戦略・標準化戦略の決定

リスクベース・ガバナンス：適正レベルを決定する6つの評価軸

1

誤りの影響 (Impact of Error)

権利喪失、侵害、契約責任、信用毀損につながるか？

2

可逆性 (Reversibility)

誤りを後から訂正・回収できるか？

3

検証可能性 (Verifiability)

一次資料、該当箇所、検索履歴から短時間で確認できるか？

4

定型性 (Routineness)

入力、判断ルール、成果物形式、例外処理が明確に定義されているか？

5

機密性 (Confidentiality)

未公開発明、係争情報、個人情報を含むか？

6

監査可能性 (Auditability)

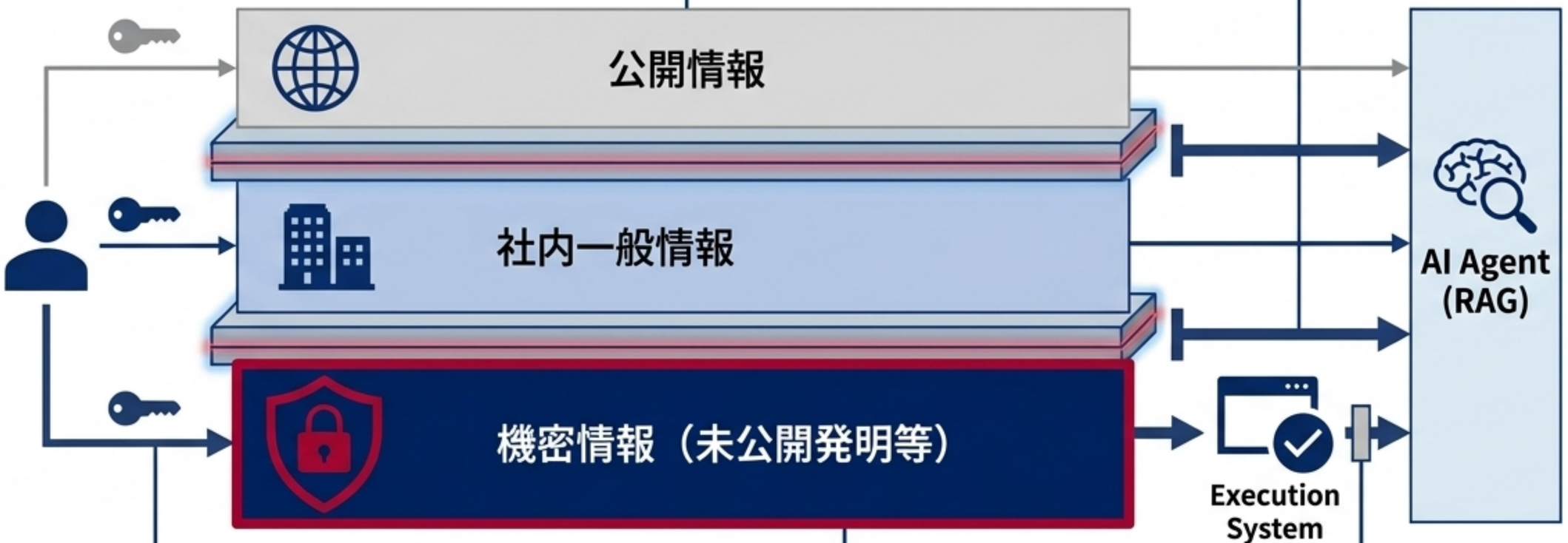
AIの入力、参照資料、出力、承認履歴を追跡できるか？

異常時に即座に停止させ、誤処理を切り戻し、人間の担当者へ引き継げる「停止・復旧性」の担保が必須。

機密情報保護のアーキテクチャ設計

層ごとの分離: 公開情報 / 社内一般情報 / 機密情報 / 特別管理情報に分け、利用可能なAI環境を厳格に定める。

RAGの権限一致: RAG（検索拡張生成）の検索権限を利用者権限と完全に一致させる。案件・共同研究先・係争単位で情報を物理的/論理的に分離する。



最小権限の原則:
AIエージェントのブラウザ、メール、ファイル、特許管理システムへのアクセス権限を最小化する。

実行系ツールの統制:
対外送信、ファイル更新、出願処理等の「実行系」機能には個別の人間承認ステップを組み込む。

評価指標：AIの出力品質とガバナンスをどう測るか

AI導入の成否は「もっともらしさ」ではなく
「重大な見落としの排除」で決まる。

品質 (Quality)

- 主要文献再現率
- 重大見落とし率
- 引用正確率
- 事実誤認率

効率 (Efficiency)

- 総工数
- 専門家の最終確認時間
- リードタイム
- 案件当たり費用

統制 (Control)

- 承認逸脱件数
- 機密情報誤入力
- 権限超過
- 監査ログ欠損
- 異常停止件数

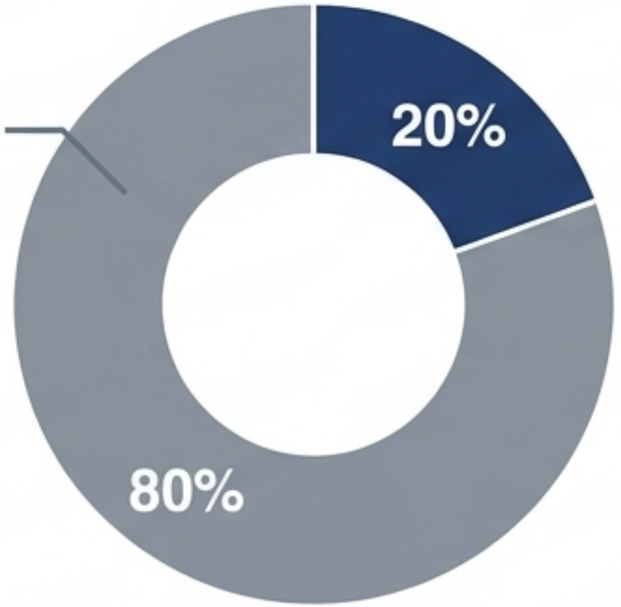
価値 (Value)

- 発明発掘件数
- 事業提案数
- リスク早期発見
- 不要権利コスト削減

段階的導入ロードマップ：ツール先行からの脱却

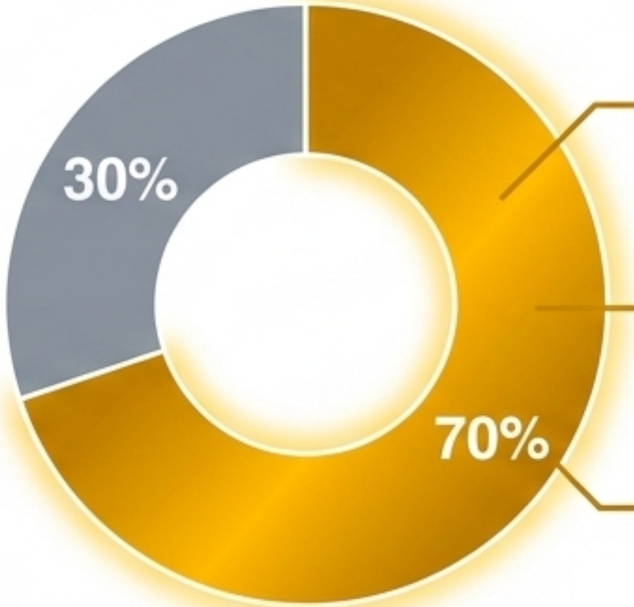
知財部門の目指す姿：「作業者」から「設計者」への進化

Before（現在）



定型的な調査、書面作成、データ整理に大半の時間を消費。

After（適応的な知財部門）



- 発明発掘と研究開発部門との共創
- 事業戦略との直接的な接続
- ライセンス・標準化交渉

知財部門の競争力を左右するのは「AIの自律度」そのものではない。
どの仕事をどこまでAIに任せ、どこで人間が責任を引き受けるか——
その「委任のアーキテクチャを設計する能力」こそが、次世代の
知財専門家に求められるコア・スキルである。