

ARC-AGI-2ベンチマークから見る、 高性能LLMが切り拓く知的財産業務の新たな地平

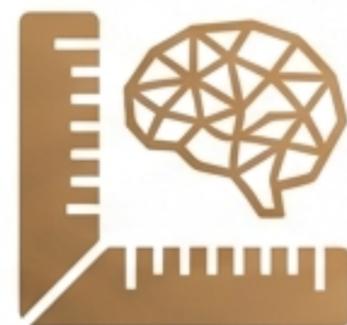
推論能力の「断絶」がもたらす業務品質の決定的な違い

エグゼクティブ・サマリー：AI活用の成否を分ける「流動性知能」



Context (背景)

知財業務へのLLM導入が進む中、モデル間の性能差が拡大している。



The Standard (基準)

ARC-AGI-2は、知識の量ではなく「未知への適応力（流動性知能）」を測る指標である。



The Impact (影響)

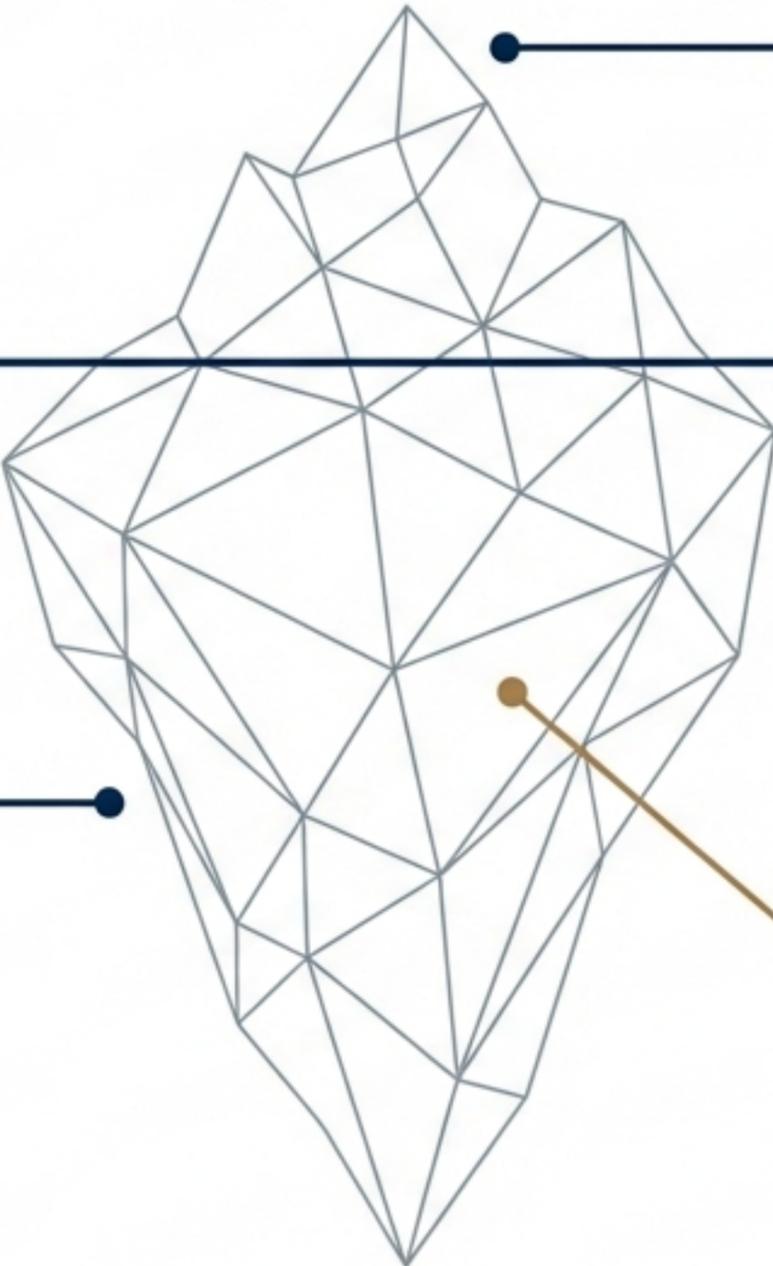
高性能モデルのみが、無効資料調査やクレーム解釈といった複雑な推論タスクを遂行できる。



The Future (展望)

人間とAIの協業（Human-in-the-Loop）が、知財専門家の能力を拡張する新たな標準となる。

一般的なLLMが抱える「推論の欠如」という課題



Low Capability Required

The diagram features a wireframe iceberg with a horizontal line representing the water surface. The small tip above the surface is labeled 'Low Capability Required', and the large submerged part is labeled 'Deep Work (High Capability)'. A callout box points to the submerged part.

- メール作成
- 単純翻訳
- 要約

Deep Work (High Capability)

- 進歩性の判断
- 均等論の解釈
- 戦略的なクレームドラフティング

Insight: 多くのLLMはこの「水面下」の業務で失敗する。なぜなら、彼らに欠けているのは知識ではなく「流動性知能」だからである。

知識の「記憶量」ではなく、未知への「適応力」を測る指標



Traditional Benchmark

Static Knowledge (静的知識)

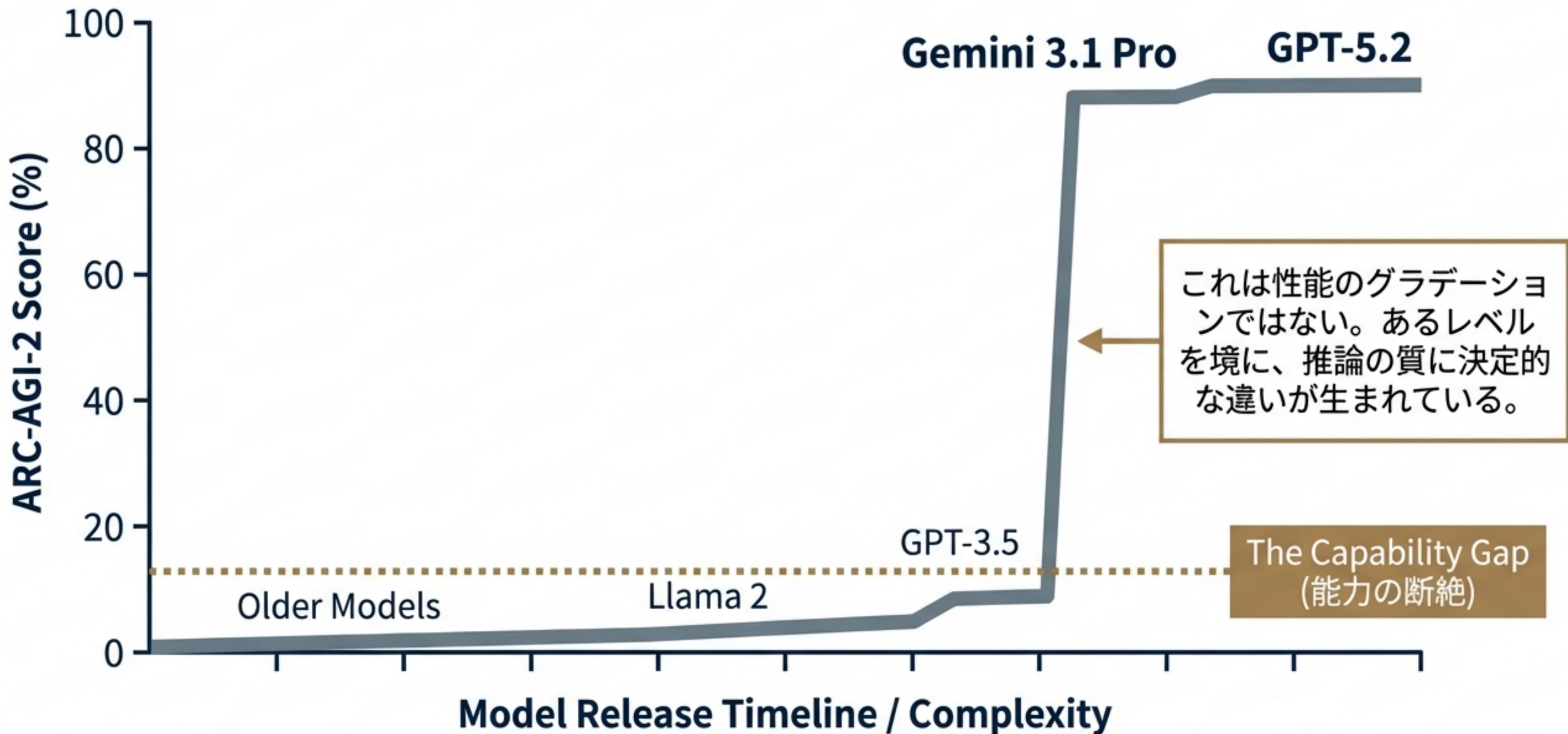


ARC-AGI-2

Fluid Intelligence (流動性知能)

ARC-AGI-2は、人間には直感的に解けるがAIには困難な「未知の課題」への適応力を測定する。これは、過去のデータの記憶ではなく、その場での推論能力を問う「AIのための司法試験」である。

明確な「能力の断絶」：最新モデルのみが到達した領域



高性能LLMと低性能LLMを分かつ「3つの推論障壁」



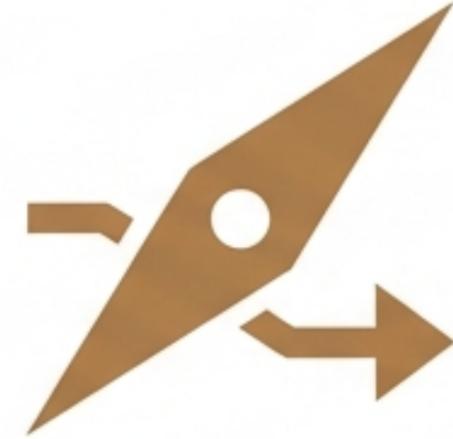
1. Symbolic Interpretation (記号的解釈)

記号を文脈の中で「意味」として捉える力。



2. Compositional Reasoning (構成的推論)

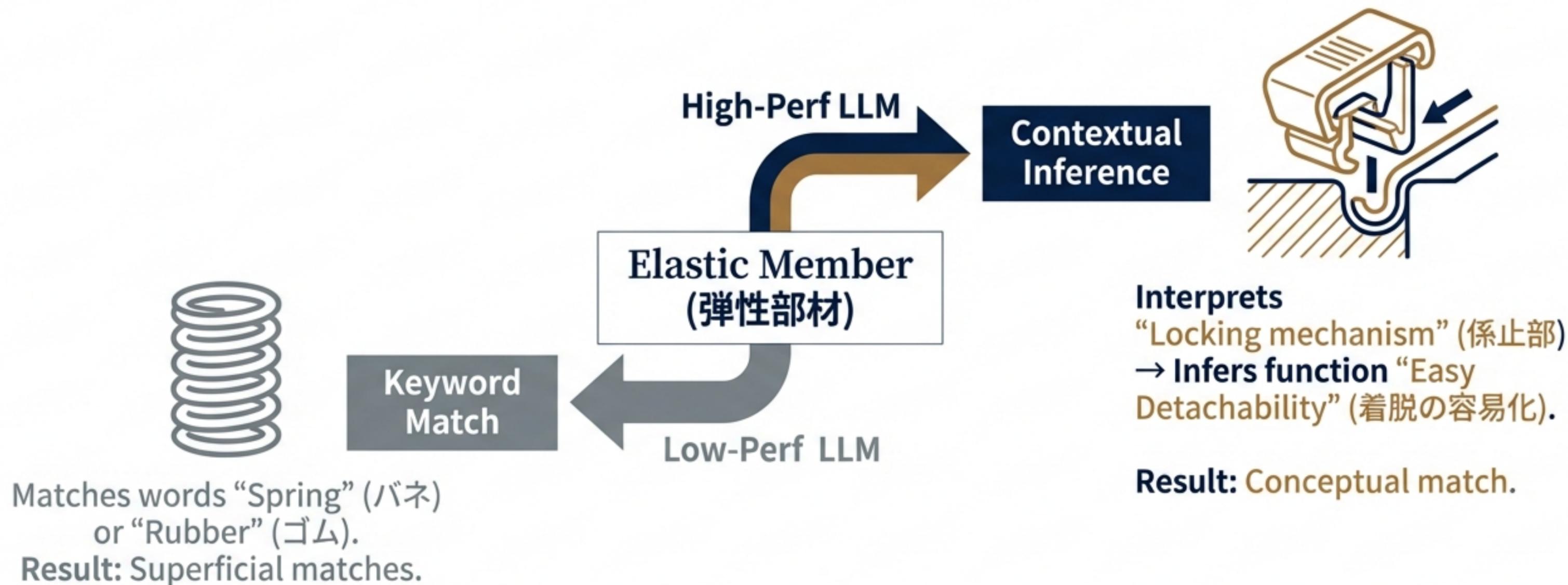
複数のルールや制約を複合的に操る力。



3. Contextual Rule Application (文脈に応じたルール適用)

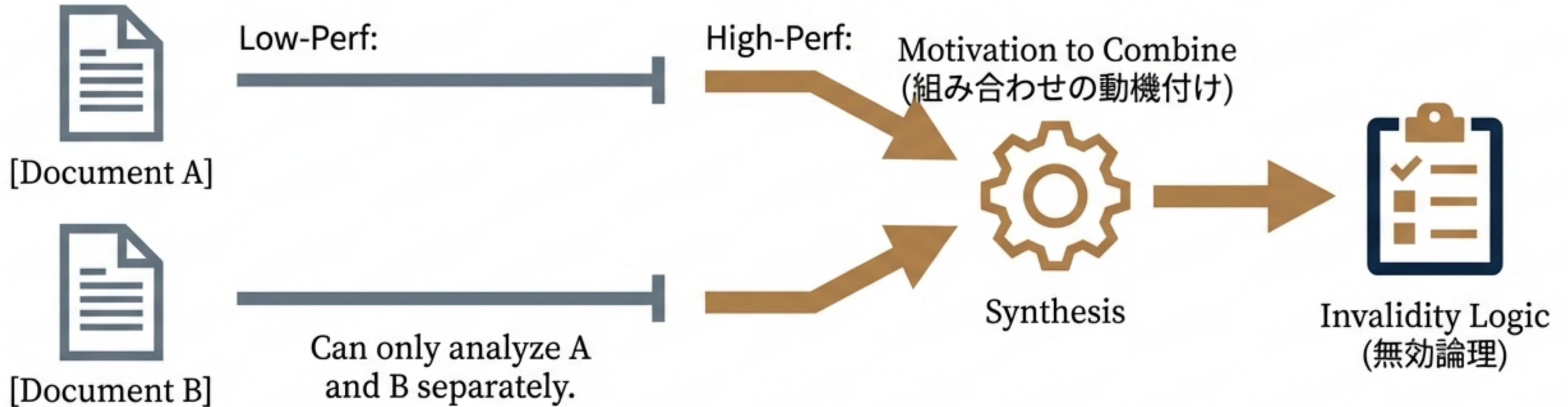
状況に応じて判断基準を動的に切り替える力。

【記号的解釈】 発明の「魂」を理解し、ノイズを排除する



Business Impact: 単語の羅列ではなく「技術的思想」を理解することで、類似用語の無関係な文献を排除し、表現が異なっても本質が同じ先行技術を発見する。

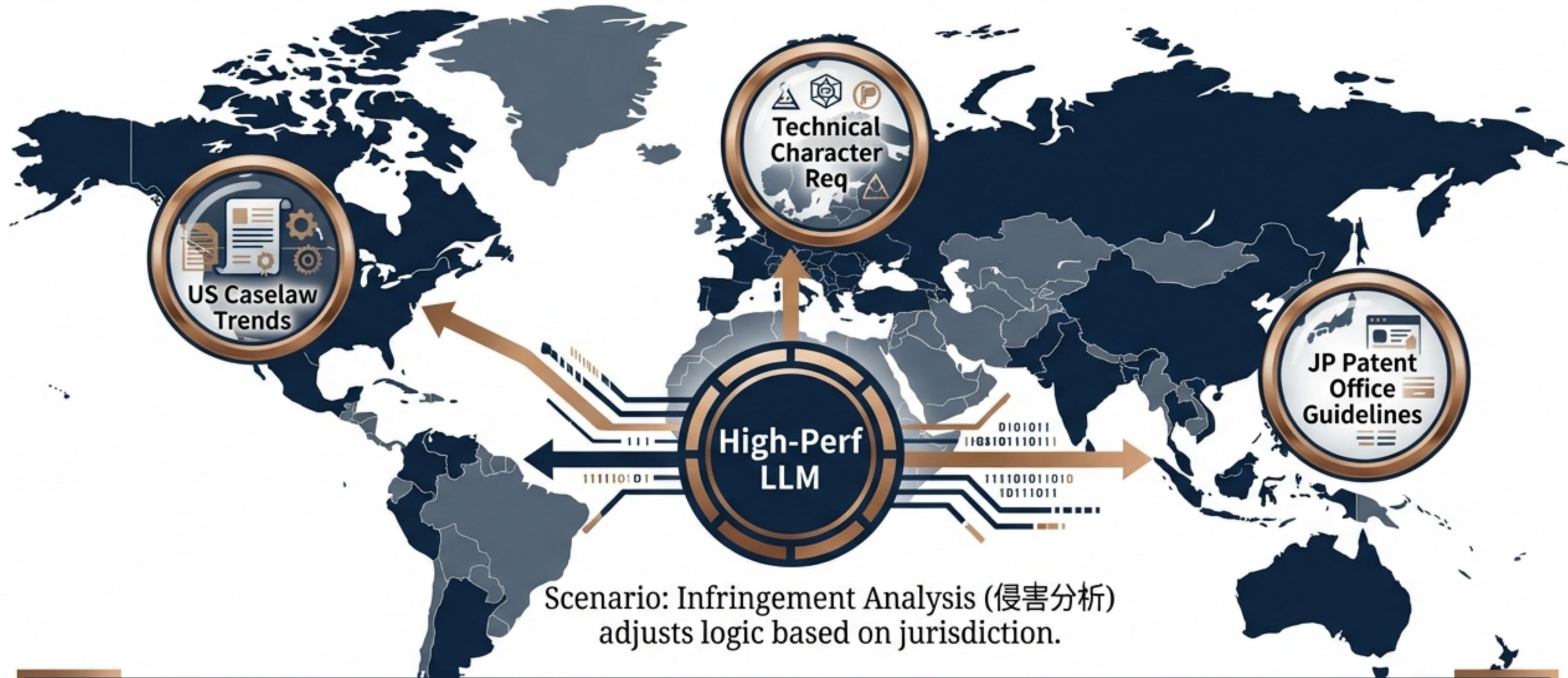
【構成的推論】複数の証拠から「進歩性欠如」の論理を構築する



Business Impact:

新規性、進歩性、記載要件など複数の法的要件を同時に満たすクレーム代替案の提案や、説得力のある無効資料調査が可能になる。

【文脈に応じたルール適用】 国や判例傾向に即した動的な判断

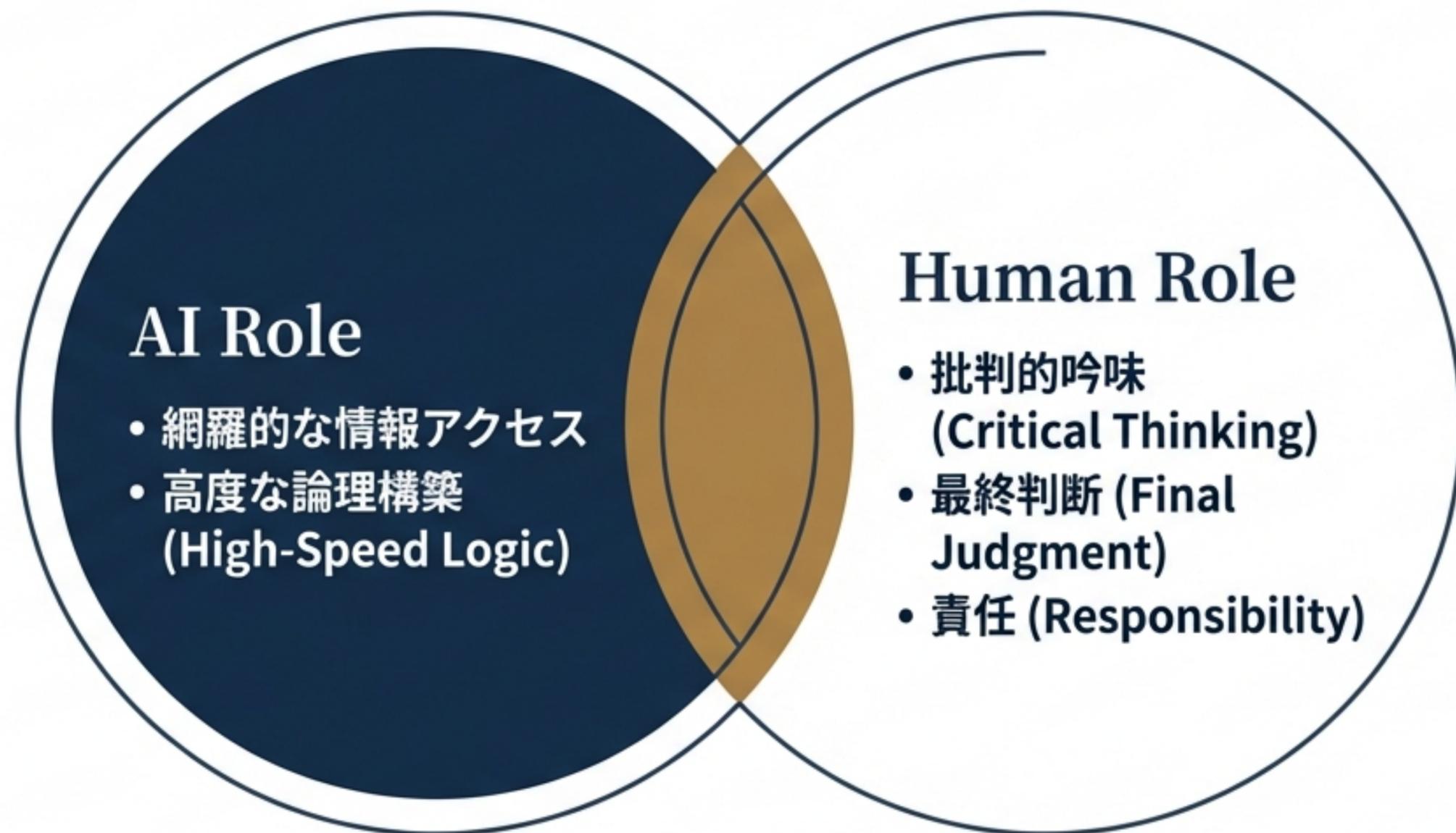


Business Impact: 画一的な回答ではなく、各国の法制度や技術分野の常識（コンテキスト）を踏まえた、精度の高いグローバル出願戦略を支援する。

比較要約：業務品質におけるリスクと価値

Task Domain	Low-Perf Risk (Risk)	High-Perf Value (Value)
Search (調査)	Keyword matching only (False Negatives/Missed Art) 	Concept matching (High Recall/Precision) 
Logic (論理)	Single-step logic 	Multi-step compositional reasoning (Strategic Invalidity Arguments) 
Strategy (戦略)	Static rules 	Dynamic context switching (Global Relevance) 

専門家の能力を拡張する「Human-in-the-Loop」



AIの出力には誤りが含まれる可能性があるため、専門家による批判的な吟味は不可欠である (Bui et al., 2025)。AIは「正解」を出す装置ではなく、専門家の思考を広げるパートナーとして機能する。

単純作業の自動化から、 戦略的な意思決定支援へ

高性能LLMは、知財専門家の
「思考のパートナー」となり得る。
経験と直感をAIの論理で補強し、
より創造的で戦略的な業務への
シフトを実現する。