

# 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

発行日: 2026年2月22日

著者: Manus AI

## 1. はじめに

近年、大規模言語モデル（LLM）の進化は目覚ましく、知的財産（IP）業務においても、先行技術調査の効率化から明細書作成の支援まで、その活用が現実のものとなりつつあります。しかし、一括りに「LLM」と言っても、その性能には大きな隔たりがあり、業務の質に決定的な違いをもたらします。本レポートでは、AIの真の推論能力を測る指標として注目される「ARC-AGI-2」ベンチマーク<sup>1</sup>を基準に、高性能LLMと低性能LLMの間に存在する「能力の断絶」を明らかにします。そして、その能力差が知的財産業務の現場で具体的にどのような違いとして現れるのか、特に高性能LLMにしかできないことに焦点を当てて分析し、未来の知財専門家とAIの協業モデルを展望します。

## 2. ARC-AGI-2が示す「推論能力の壁」

ARC-AGI-2（Abstraction and Reasoning Corpus for Artificial General Intelligence 2）は、知識の記憶量ではなく、未知の課題に直面した際の適応能力、すなわち「流動性知能」を測定するために設計されたベンチマークです<sup>2</sup>。人間にとっては直感的に解ける問題群である一方、現在の多くのAIにとっては極めて困難な課題とされています。

最新のARC-AGI-2リーダーボード<sup>3</sup>を見ると、その困難さが浮き彫りになります。Gemini 3.1 ProやGPT-5.2といった最先端の高性能LLMがようやく人間には及ばないながらも有意なスコアを記録し始めた一方で、少し前のモデルや性能の低いモデルのスコアは軒並み0%に近いままです。これは、単なる計算能力や知識量の差ではなく、AIの「推論の質」において、あるレベルを境に**明確な能力の断絶**が存在することを示唆しています。

ARC-AGI-2の技術レポート<sup>4</sup>によれば、このベンチマークは特に以下の3つの高度な推論能力を評価するように設計されており、これらこそが高性能LLMと低性能LLMを分ける「壁」となっています。

- 記号的解釈 (Symbolic Interpretation):** 記号を単なる視覚パターンとしてではなく、文脈の中で与えられた意味を持つものとして解釈する能力。
- 構成的推論 (Compositional Reasoning):** 複数のルールや制約を同時に、あるいはそれらの相互作用を考慮しながら適用し、結論を導き出す能力。
- 文脈に応じたルール適用 (Contextual Rule Application):** 周囲の状況や文脈によって、適用すべきルールを動的に切り替える能力。

知的財産業務は、まさにこれらの高度な推論能力が求められる領域です。特許文献の解釈は、単語の定義だけでなく、明細書全体の文脈、技術分野の常識、そして法律や判例といった複数のルールを複合的に適用して行われます。次章では、この3つの能力を軸に、高性能LLMだけが遂行可能な知的財産業務を具体的にみていきます。

### 3. 高性能LLMにしかできない知的財産業務 — 3つの能力に基づく分析

ARC-AGI-2が要求する3つの推論能力は、知的財産業務における専門家の思考プロセスと深く結びついています。この能力の有無によって、LLMが提供できる価値は「単純作業の補助」から「高度な意思決定支援」へと大きく変化します。

推論能力	知的財産業務におけるタスク例	高性能LLMにしかできないこと	低性能LLMの限界
記号的解釈	クレーム解釈、発明の要旨認定、先行技術の評価	<b>発明の「魂」を理解する:</b> クレームの構成要件の組み合わせが、どのような技術的課題を解決するための「発明」という一つの意味あるコンセプトを形成しているかを理解する。例えば、「弾性部材」と「係止部」が「着脱の容易化」という機能的意味を持つことを解釈する。	<b>単語のリストとして処理:</b> クレームを単語の集合としてしか認識できず、キーワードのマッチングに終始する。用語が同じでも技術的思想が異なる文献を誤って関連ありと判断する。
構成的推論	無効資料調査、戦略的なクレーム作成支援、侵害回避設計	<b>複数の証拠から論理を構築する:</b> 複数の先行技術文献を組み合わせ、対象特許の進歩性を否定する論理（組み合わせの容易想到性）を構築する。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を複数提案する。	<b>一度に一つのルールしか適用できない:</b> 新規性についてはコメントできても、進歩性との関係性を考慮した分析はできない。複数の文献を提示されても、それらを戦略的に組み合わせる推論は不可能。
文脈に応じたルール適用	侵害成否の判断、各国法制度を考慮したグローバル出願戦略の立案	<b>状況に応じて判断基準を切り替える:</b> クレーム中の用語が、明細書の記載や技術分野の常識、判例に基づき、ど	<b>常に同じルールを画一的に適用:</b> 文脈を無視し、ある国で認められた判例を法制度が異なる国にもそのまま適用

のように解釈されるべきかを判断する。例えば、同じ「通信」という言葉でも、国や技術分野の文脈で解釈の範囲が変わることを理解し、侵害分析の精度を高める。

できるかのような誤った分析を行うリスクがある。

## 【能力1】記号的解釈：発明の本質を見抜く

特許実務において、クレーム（特許請求の範囲）に記載された用語は、辞書的な意味だけでなく、明細書全体の文脈の中で定義された「技術的な記号」として機能します。高性能LLMは、この記号が持つ発明の核心的な意味、すなわち「発明の要旨」を捉えることができます。これにより、単なるキーワード検索では見つけられない、表現は異なるが本質的に同じ技術思想を持つ先行技術文献を発見したり、逆に、用語は似ていても発明の核心が異なる文献をノイズとして排除したりすることが可能になります。これは、調査の質を飛躍的に向上させるだけでなく、侵害訴訟におけるクレーム解釈の妥当性を検討する上でも極めて重要です。

## 【能力2】構成的推論：複数の制約から最適解を導く

無効資料調査は、単一の文献で新規性を否定するだけでなく、複数の文献を組み合わせることで進歩性の欠如を主張することが頻繁に行われます。高性能LLMは、複数の文献の内容をそれぞれ理解した上で、それらを組み合わせる「動機付け」が当該技術分野に存在したか、組み合わせによって予期せぬ効果（顕著な効果）が生じないか、といった複数の視点を複合的に評価し、説得力のある無効論理の仮説を構築することができます。これは、低性能LLMが個々の文献を評価することしかできないのとは対照的です。この能力は、攻撃だけでなく、自社の特許網を強化するための戦略的なクレーム作成支援にも応用できます。

## 【能力3】文脈に応じたルール適用：状況に応じて判断を切り替える

特許権の効力は、国や地域によって異なります。また、同じ法律の条文でも、技術分野や時代背景によってその解釈が変化することがあります。高性能LLMは、このような「文脈」を理解し、例えば「米国のソフトウェア特許訴訟における近年の判例傾向」といった特定のコンテキストに基づいた侵害リスクの分析や、「欧州と中国で権利化しやすいクレーム表現」といった、グローバルな出願戦略の立案支援が可能になります。低性能LLMが文脈を無視して画一的な回答しかできないのに対し、高性能LLMは状況に応じた柔軟な思考を提供し、より精度の高い意思決定をサポートします。

## 4. 結論：専門家を拡張する「思考のパートナー」としてのLLM

ARC-AGI-2ベンチマークが示すように、高性能LLMと低性能LLMの間には、単なる性能のグラデーションではなく、高度な「推論能力」の有無という明確な断絶が存在します。知的財産という、深い洞察と多角的な分析が求められる専門性の高い業務において、この差は決定的な価値の違いを生み出します。

低性能なLLMが単純作業の自動化に留まるのに対し、**高性能LLMは、知財専門家の「思考のパートナー」として、その能力を拡張する存在**になり得ます。専門家が持つ経験や直感を、LLMが持つ網羅的な情報アクセス能力と高度な論理構築能力で補強し、これまで以上に質の高い戦略的な意思決定を迅速に行う。これが、高性能LLMが拓く未来の知的財産業務の姿です。

もちろん、現状の高性能LLMですら完璧ではなく、その出力には誤りが含まれる可能性があります。だからこそ、最終的な判断は人間が行うべきであり、AIの出力を鵜呑みにせず、批判的に吟味し、修正・活用する専門家の役割、すなわち「**Human-in-the-Loop (HITL)**」<sup>5</sup>の重要性は、今後ますます高まっていくでしょう。高性能LLMの登場は、専門家の仕事を奪うのではなく、より創造的で戦略的な業務へとシフトさせる契機となるのです。

---

## 参考文献

- [1] ARC Prize. "ARC-AGI-2".
- [2] Chollet, F. (2019). "On the Measure of Intelligence". arXiv:1911.01547.
- [3] ARC Prize. "Leaderboard".
- [4] Chollet, F., et al. (2026). "ARC-AGI-2: A New Challenge for Frontier AI Reasoning Systems". arXiv:2505.11831v2.
- [5] Bui, L. V., et al. (2025). "Advancing patent law with generative AI: Human-in-the-loop systems for AI-assisted drafting, prior art search, and multimodal IP protection". World Patent Information, 80, 102341.