

知財インテリジェンスの 地殻変動

PatentBenchとAIエージェント「Eureka」
が証明した2026年の実務革新

Based on the 2026 PatentBench Report &
Inter Regular & Noto Serif JP Regular



イノベーションの加速 と「調査の壁」

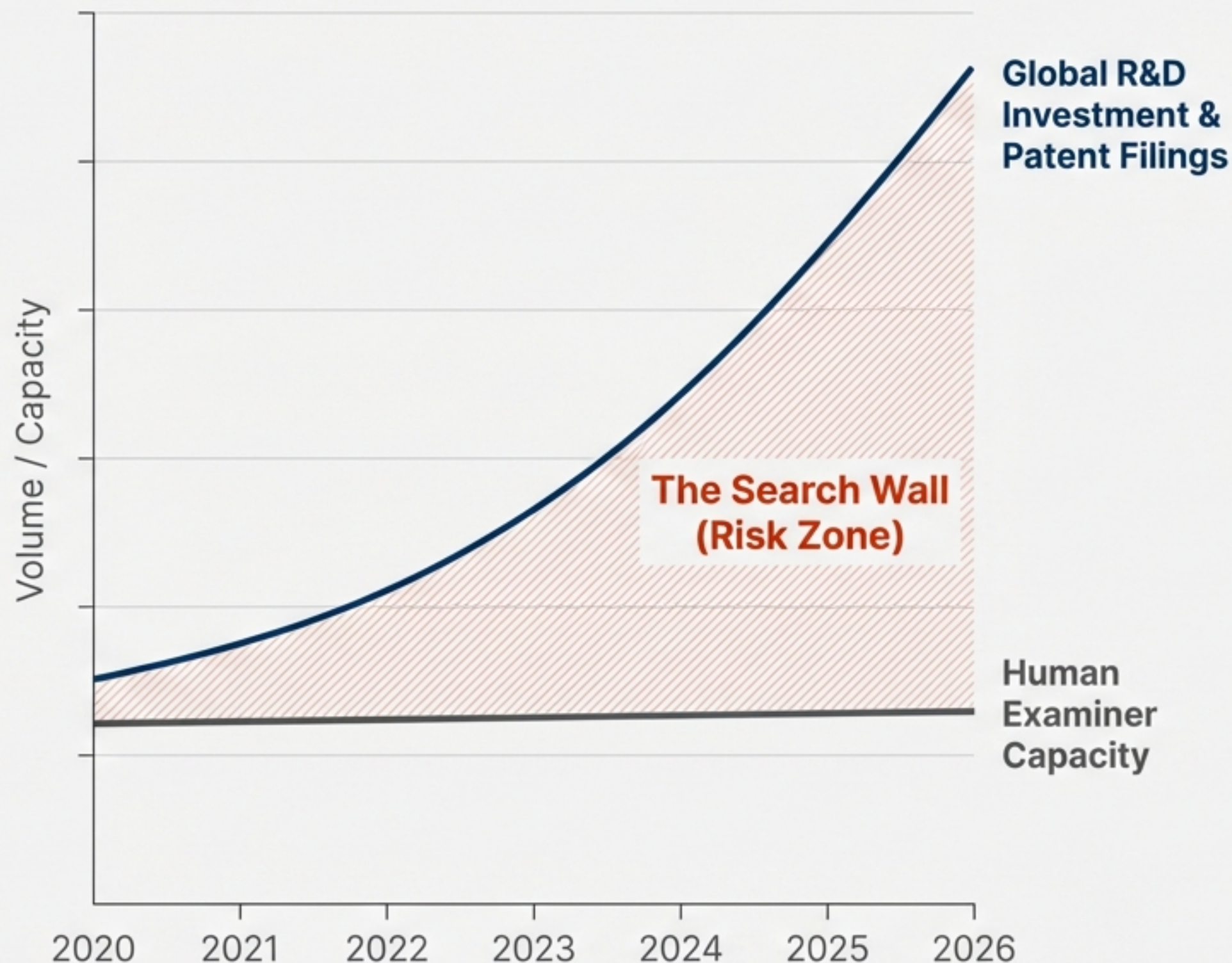
The Bottleneck

イノベーションを保護するIPシステムは依然として労働集約的である。特許庁審査官や知財専門家のリソースは限界に達している。

The Cost

新規性確認やFTO調査には、熟練者でも数日から数週間を要する。この遅れはR&Dスピードを阻害するだけでなく、見落としによる無効リスクや侵害訴訟という「潜在的な爆弾」を生む。

The Widening Gap



汎用LLMへの熱狂と、現場の幻滅

なぜChatGPTは「銀の弾丸」にならなかったのか

The Promise



流暢な文章生成と
一般的推論能力

The Reality in IP

Semantic Mismatch

汎用LLMは「ネジ」と「締結部材」、「バネ」と「弾性体」のような特許特有の上位概念化・抽象化を理解できない。



Hallucinations

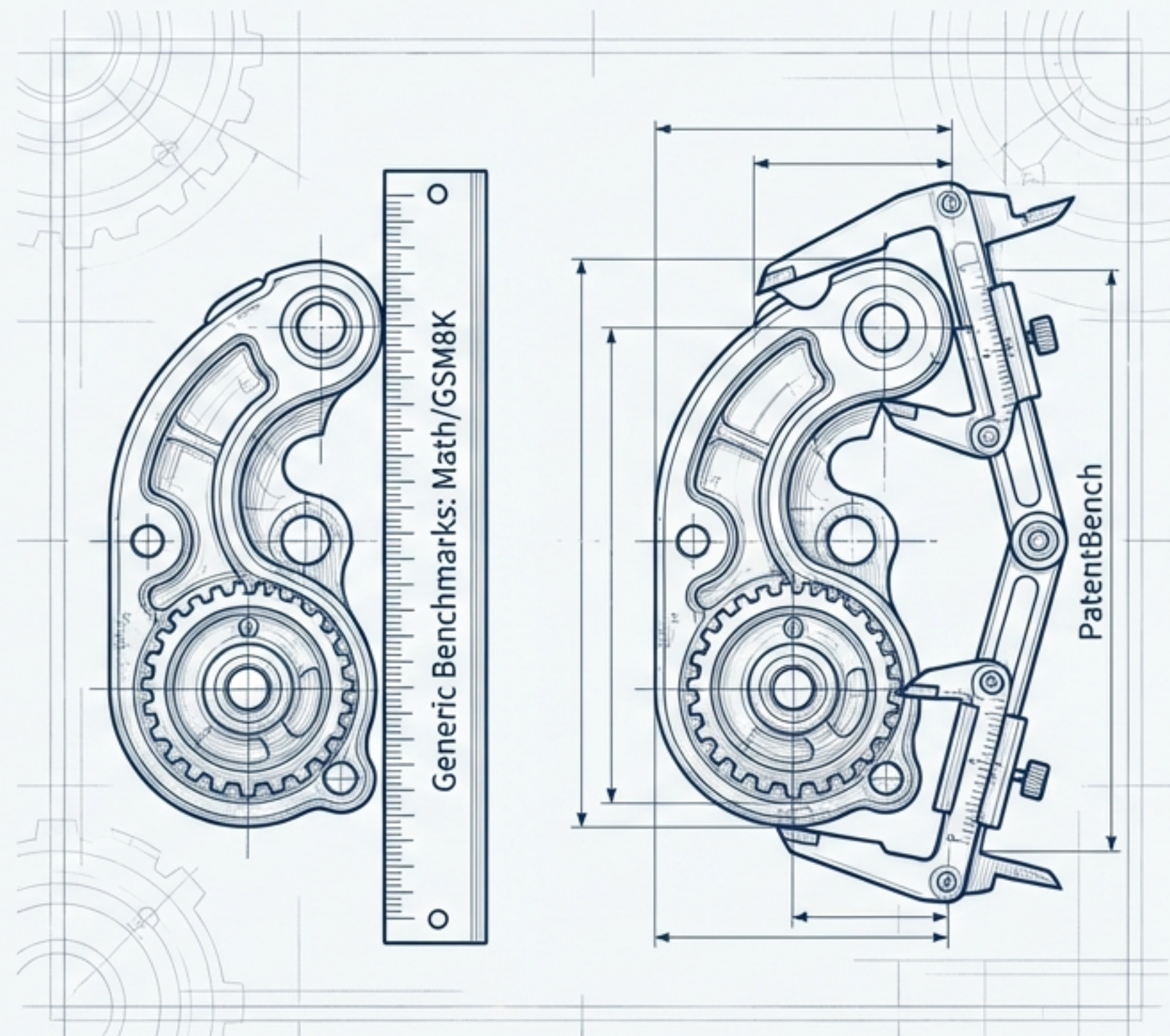
存在しない特許番号や判例の捏造。
法的正確性が生命線の業務において致命的な欠陥。



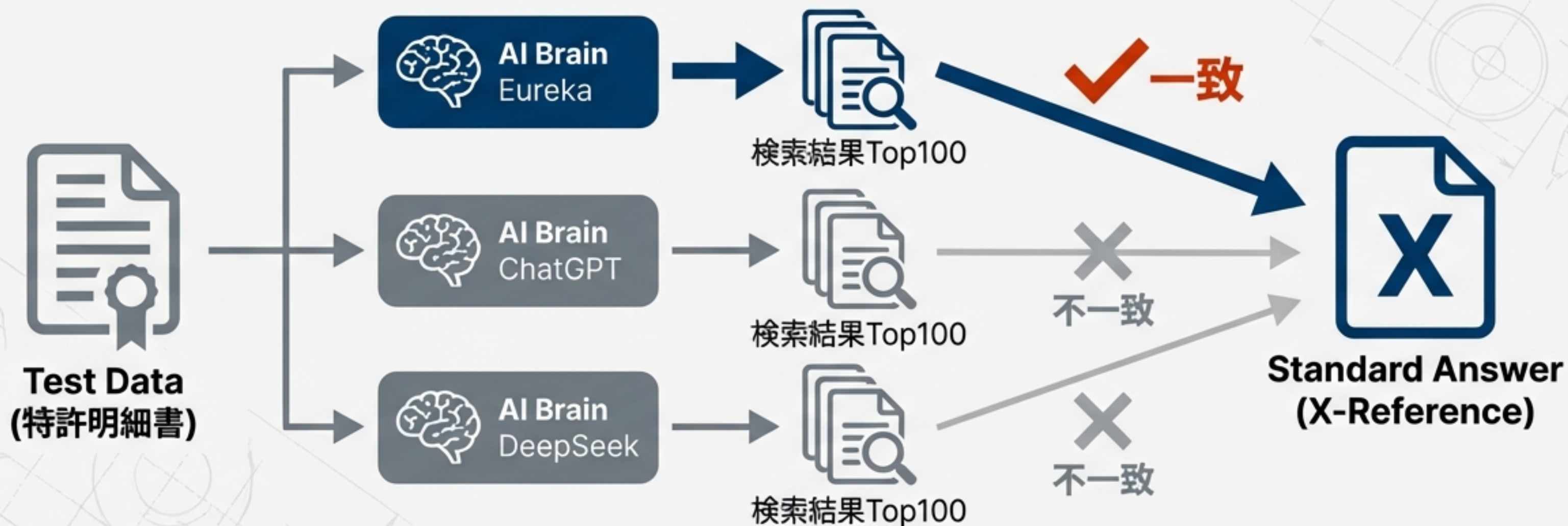
結論：一般的な推論能力は、専門的な知財コンピテンシーとイコールではない。

信頼を取り戻すための「新しいものさし」：PatentBench

- MMLU（一般的知識）やGSM8K（数学）は、特許の新規性判断や請求項解釈の能力を測るには不適切である。
- Patsnapが構築した業界初のベンチマーク「**PatentBench**」。
- **言語生成能力ではなく、「実務遂行能力」を測定する。**
- 基礎スキル（Q&A、翻訳）に加え、応用能力（新規性調査、FTO分析）を客観的に数値化する。



評価メカニズム：審査官引用を「正解」とする検証フロー



Test Data

US, CN, EP, WOの主要特許庁で審査された340件の特許ファミリー（英語68%、中国語32%）。

The Ground Truth

審査官が拒絶理由通知で引用した「X文献（新規性否定文献）」を正解と定義。

The Challenge

AIは明細書テキストのみから、膨大なデータベースを検索し、審査官が見つけたのと同じ決定的な文献を発見できるか？

実務的価値を測る2つの指標

1. X Hit Rate (Xヒット率)

The Sniffer Dog (目利き力)



定義: 検索結果の上位100件に、少なくとも1件の正解 (X文献) が含まれる割合。

価値: 初期スクリーニングの効率性。「この発明は新規性がない」という証拠を早期に掴み、無駄な投資を抑制する能力。

2. X Recall Rate (Xリコール率)

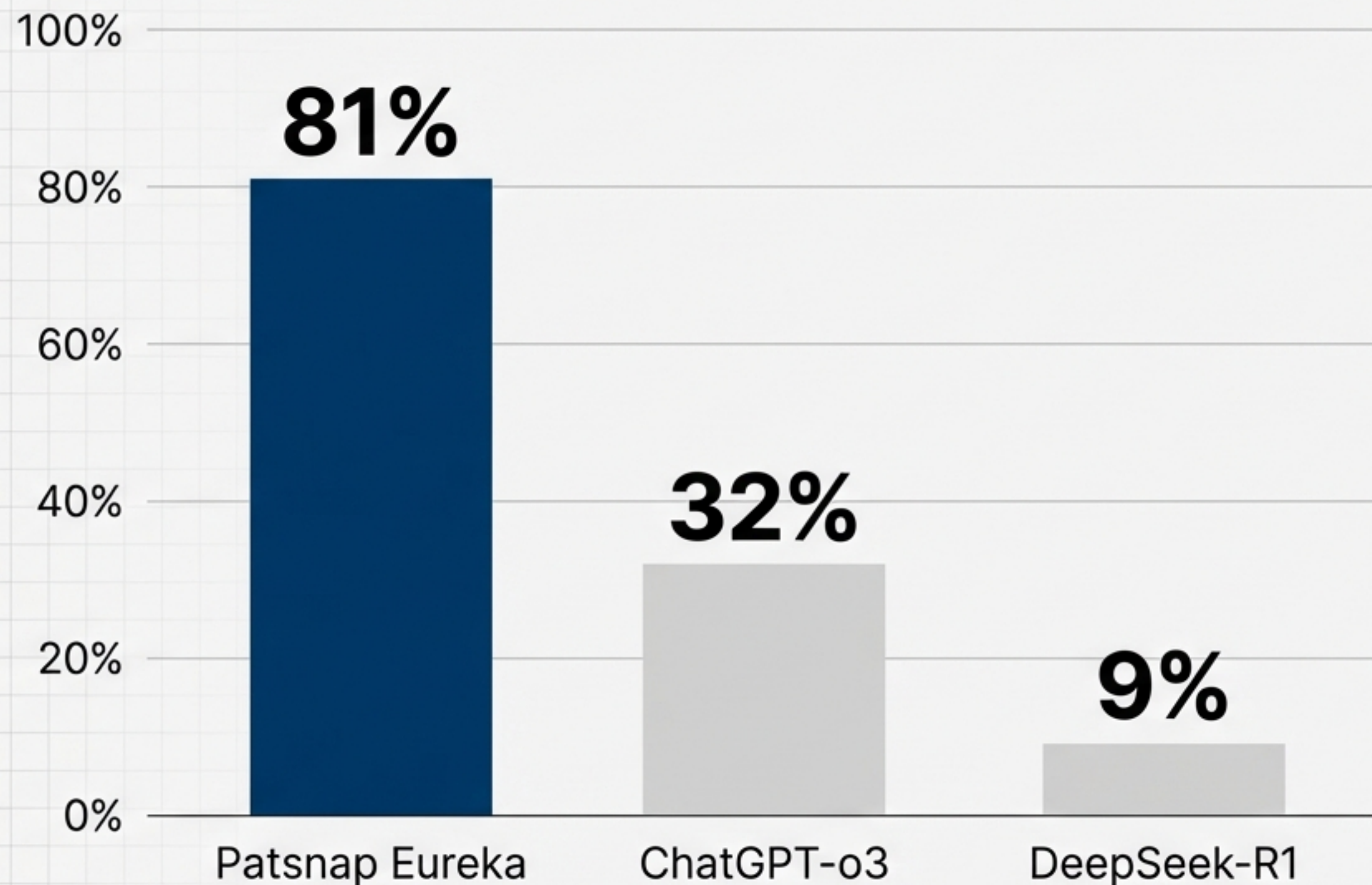
The Safety Net (網羅性)



定義: 全正解X文献のうち、AIが何割を拾い上げることができたか。

価値: FTOや無効資料調査における信頼性。リスク要因 (他社特許) をいかに漏れなく洗い出せるかを示す。

決定的な性能格差：専用AI vs 汎用LLM



汎用的な推論能力 (DeepSeek) だけでは、特許調査の複雑な検索には全く対応できないことが証明された。

Patsnap Eureka: 81%.
5回中4回のケースで審査官レベルの引用文献を自力で発見。

なぜDeepSeek-R1は「9%」に沈んだのか？



1. Domain Knowledge Void (ドメイン知識の欠如)

強化学習で論理能力は高いが、特許分類（IPC/CPC）、審査基準、公報構造の知識が学習データに含まれていない。



2. The Web Search Wall (Web検索の限界)

表面的な技術ブログしか参照できず、特許文脈でのクロスリンガル検索（構造化された対訳データ）を持たないため、多言語検索で失敗した。



2. Baixing Search Wall (言語の壁)

強化学習技術ブログしながら、特許文脈（対許リンガル（構造化された知識が学れず対訳データ）を）に含まれていない。



3. Language Barrier (言語の壁)

中国発モデルでありながら、特許文脈でのクロスリンガル検索（構造化された対訳データ）を持たないため、多言語検索で失敗した。

Xリコール率「36%」の解釈：AIはパイロットか、ナビゲーターか

36%

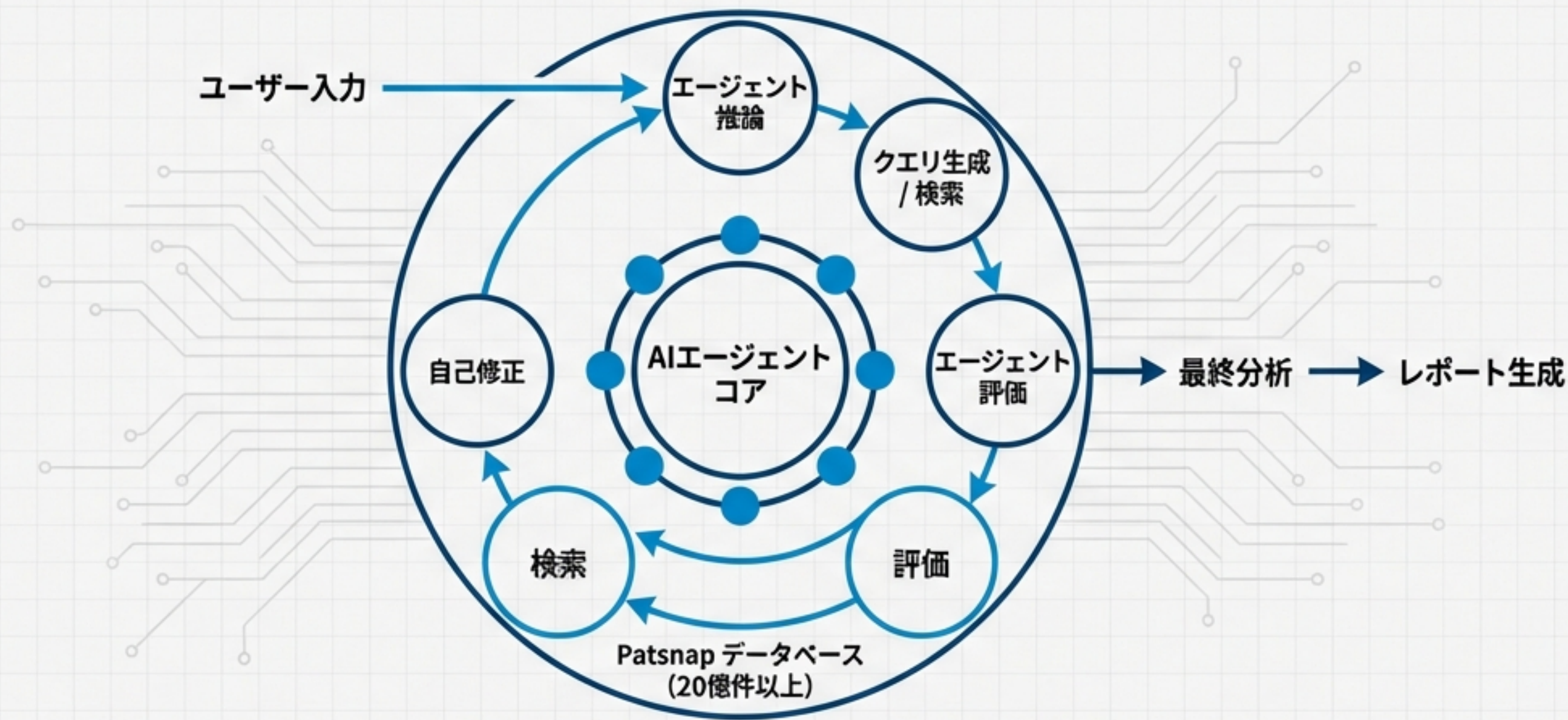
Eureka Total Recall Rate

The Reframe:

- **Impossible Standard:** 熟練調査員でも短時間で全てのX文献を網羅するのは至難。
- **Productivity Multiplier:** Top 100という確認可能な範囲に、重要文献の1/3以上を提示できる意味。
- **ROI:** 「3日間の反復的な検索作業」→「3時間の高品質な分析作業」。

AIは完全な代替（Pilot）ではなく、初動負荷を劇的に下げる強力な支援者（Navigator）である。

Eurekaの頭脳：RAGを超えた「RAT」アーキテクチャ



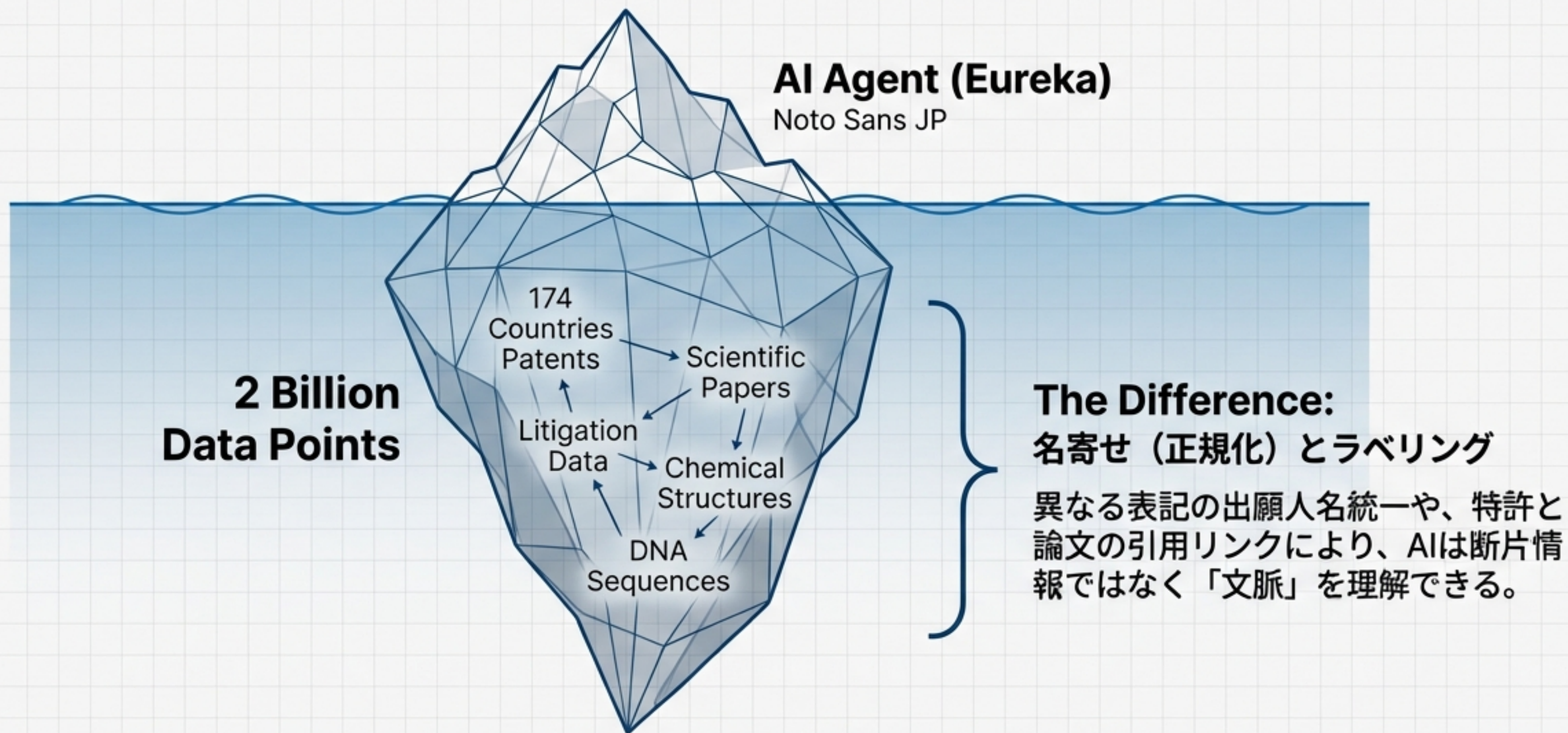
RAG (Retrieval-Augmented Generation)

最新の公報データに基づきハルシネーションを抑制。

RAT (Retrieval-Augmented Thinking)

熟練サーチャーの思考プロセスを模倣。検索 → 結果評価 → クエリ修正 → 再検索というループを自律的に実行。

知能の源泉：20億件の「構造化された知」



AIは完全な代替（Pilot）ではなく、初動負荷を劇的に下げる強力な支援者（Navigator）である。

Patsnap Eureka

エージェント・エコシステム：全方位の実務支援

Novelty Search

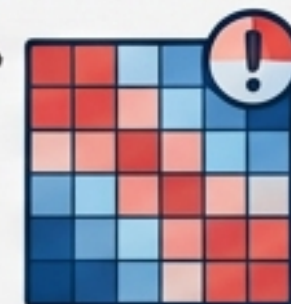
Inter Regular

構成要件分解、対比表（クレーム
チャート）作成。数日を数時間に。



FTO Search

製品仕様と他社クレーム
のヒートマップ分析。
リスク可視化。



Patent Drafting

発明届出書から明細書ド
ラフト(背景、実施例、ク
レーム)を自動生成。



Tech Scout

課題入力から解決策を
持つ特許・企業・大学
を発見。



Patent Translation

言語の壁を取り払う
高精度翻訳。



先行企業の活用事例とROI

Canon Inc.

Slate Grey
Inter Regular

アイデアから実行までのイノベーションサイクルを加速。「より速く、よりスマートな意思決定」を実現。



National University of Singapore (NUS)

Slate Grey
Inter Regular

数万ドル規模の人件費・データ費を削減。AIインサイトが技術移転 (Tech Transfer) の共通言語として機能。



Grab

Slate Grey
Inter Regular

競合ポートフォリオ分析と自社技術評価に活用。ビジネス戦略と知財戦略を統合。



2026年の知財マネジメント：戦略家への転換

Role Evolution



Before

「検索」と「読解」
(Manual Search & Reading)



After

「判断」と「戦略策定」
(Judgment & Strategy)

AIが見つけた文献の無効性判断や
回避策検討に集中する。

Governance & Security



- **Warning:** 無料Web版LLMへの未公開発明入力は機密漏洩リスク（厳禁）。
- **Enterprise Standard:** ISO 27001 / SOC 2 Type 1認証。顧客データは学習に利用されない保証。

結論：AIとの共創が拓く未来

Capability: Hit Rate 81%は、AIが実務に耐えうるパートナーであることを証明した。

Reality: Recall Rate 36%は、人間の監督（Supervision）が不可欠であることを示す。

Yes, but with supervision

定型業務をAIに任せ、人間はイノベーションの価値を最大化する創造的業務に注力せよ。それがEurekaの先にある新しい知財実務である。