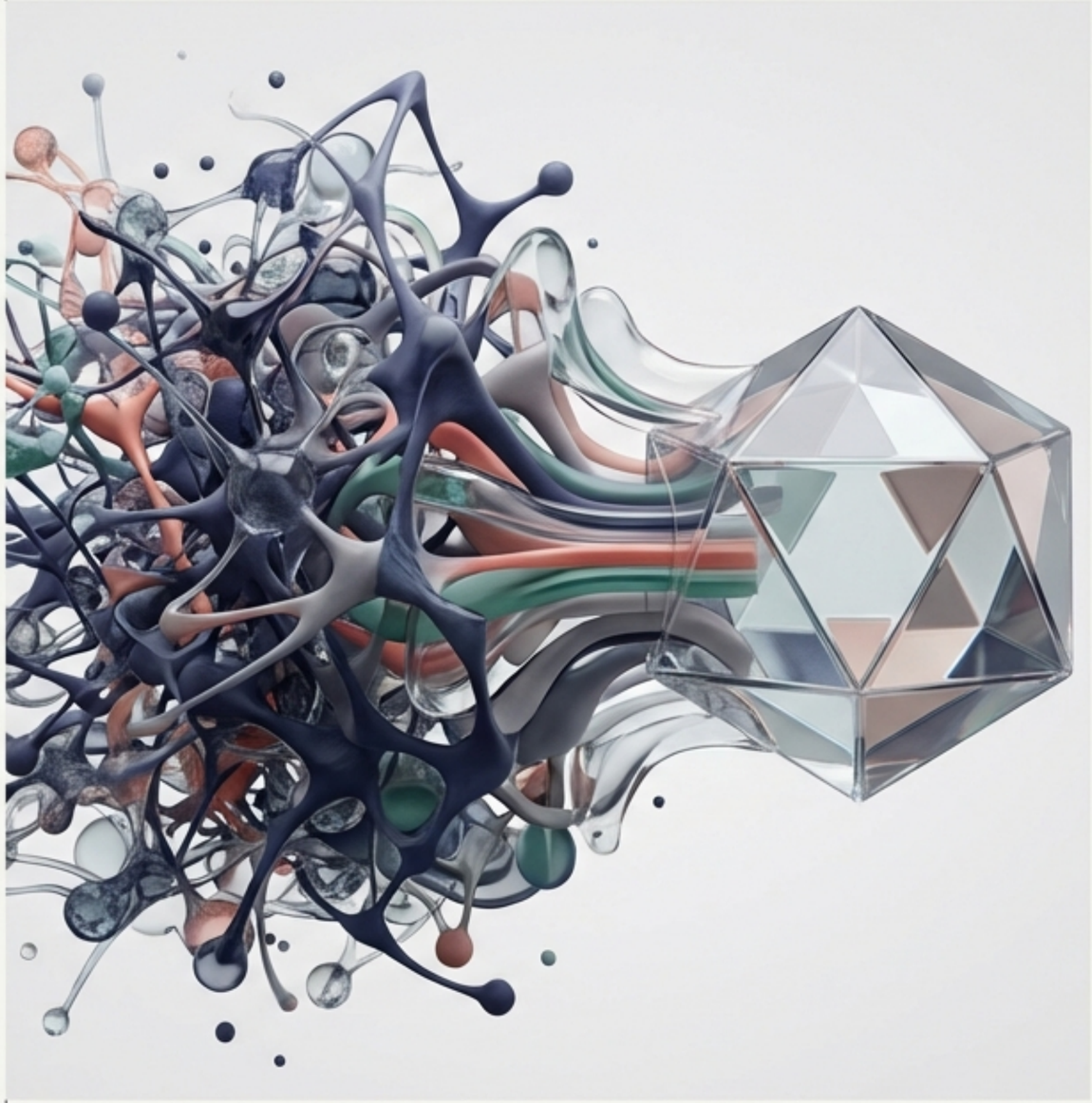
日本の主要AIモデル：高度推論・数理科学的適応性調査報告書（2026年版）

System 1からSystem 2へ ——「思考するAI」への進化と各社の戦略的ポジション

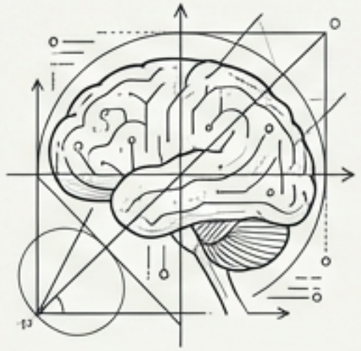
対象企業：ELYZA, Preferred Networks, SoftBank, NTT, NEC, CyberAgent

本資料は、急速に進展する「推論モデル（Reasoning Models）」の世界的潮流において、日本の主要AIプレイヤーがどのように適応し、数学・物理・科学的発見の領域で競争力を発揮しているかを分析するものである。



結論：日本のAIは推論能力を獲得したが、その戦略は3方向に分化している

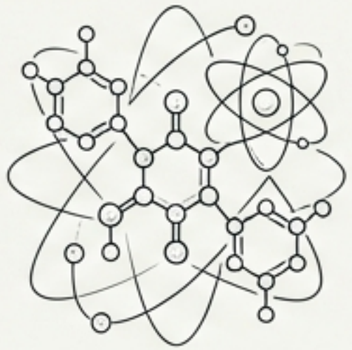
問い：「日本のモデルは数学や物理などに秀でている推論モデルに対応していますか？」
回答：「はい。ただし各社の『強み』は明確に異なる領域へ特化しています。」



汎用数学・論理
(General Math/Logic)

Key Player:
Main Player: ELYZA

OpenAI o1に匹敵する「思考プロセス（CoT）」の実装。数学的証明や論理パズルでの性能を追求。



物理シミュレーション
(Physical Simulation)

Key Player:
Main Player: Preferred Networks (PFN)

言語解説にとどまらず、原子レベルの物理法則そのものを高速計算する「Matlantis」を展開。

産業・インフラ基盤
(Industrial Application)

Key Players:
Main Players: SoftBank / NTT / NEC

合成データによる基礎体力向上（SB）、視覚情報の読解（NTT）、論理的信頼性の担保（NEC）による実社会適応。

生成AIのパラダイムシフト：「流暢な会話」から「熟考する知性」へ

System 1（直感・速度）

- 従来のLLM
- 即座に回答
- 確率的な流暢さを重視
- 苦手：複雑な論理、数学の証明

Evolution

System 2（熟考・論理）

- 推論モデル (Reasoning Models)
- 「思考の連鎖 (CoT)」による自己検証
- Test-Time Compute (思考時間)の確保
- 得意：難解な数学、物理法則、コーディング

AIの役割は「チャットボット」から「科学的発見（Scientific Discovery）の主体」へと変化している。

2026年現在の推論モデルに求められる3つの技術要件

01

テストタイム・コンピュート（Test-Time Compute）

回答出力前に計算リソースを費やして「思考」する時間を確保し、
内部で解法探索と自己修正を行うプロセスの有無。

02

強化学習（RL）による思考プロセスの最適化

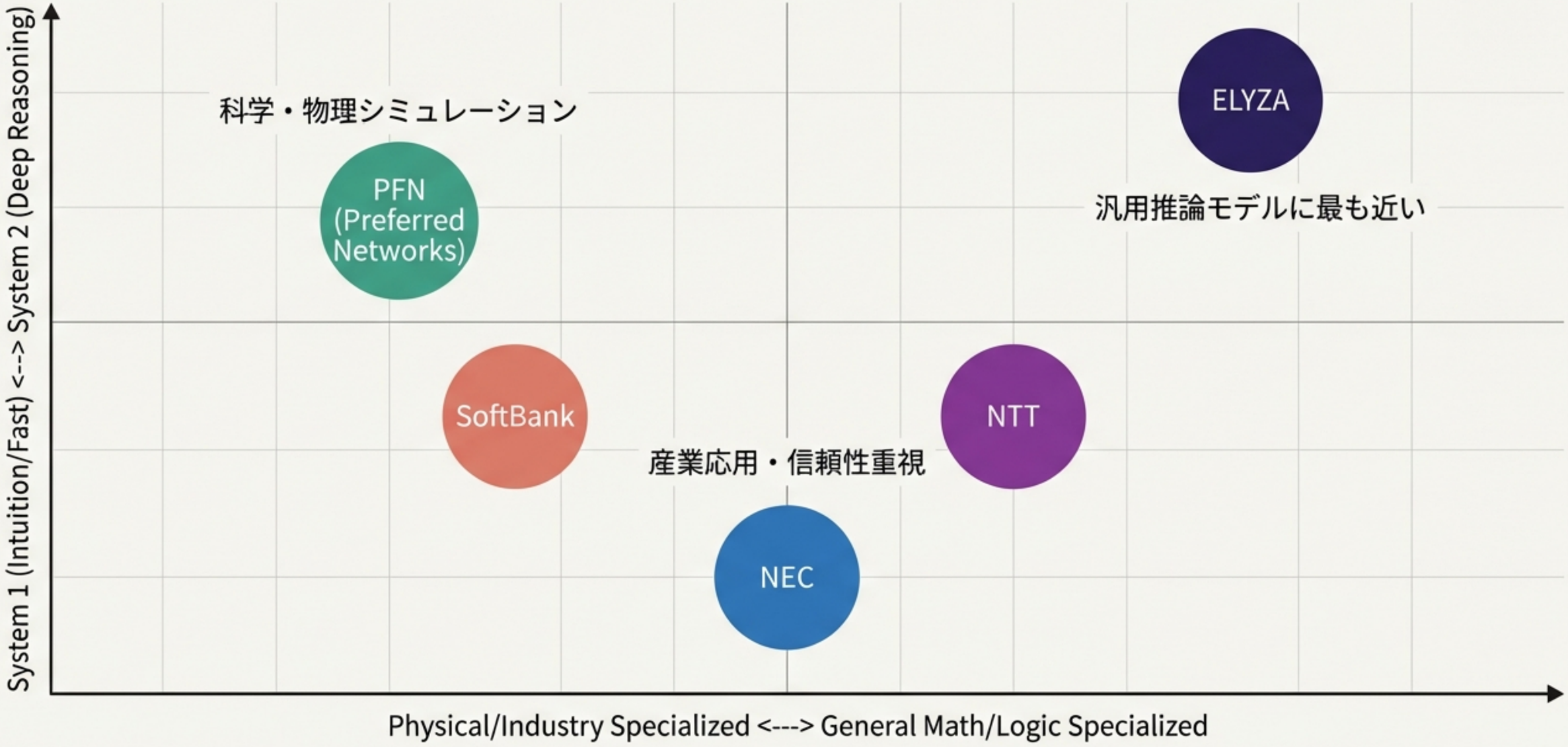
正解に至る「考え方」自体を報酬として学習（Post-training）。
未知の問題へのアプローチ能力を評価。

03

評価ベンチマーク（Benchmarks）

- MATH / MATH-500: 高度な数学的推論能力
- GPQA: 博士号レベルの科学的難問（Google検索不能）
- FrontierMath: 既存モデルでは解けない研究レベルの数学問題

戦略的ポジショニングマップ：推論の「深さ」と「対象領域」



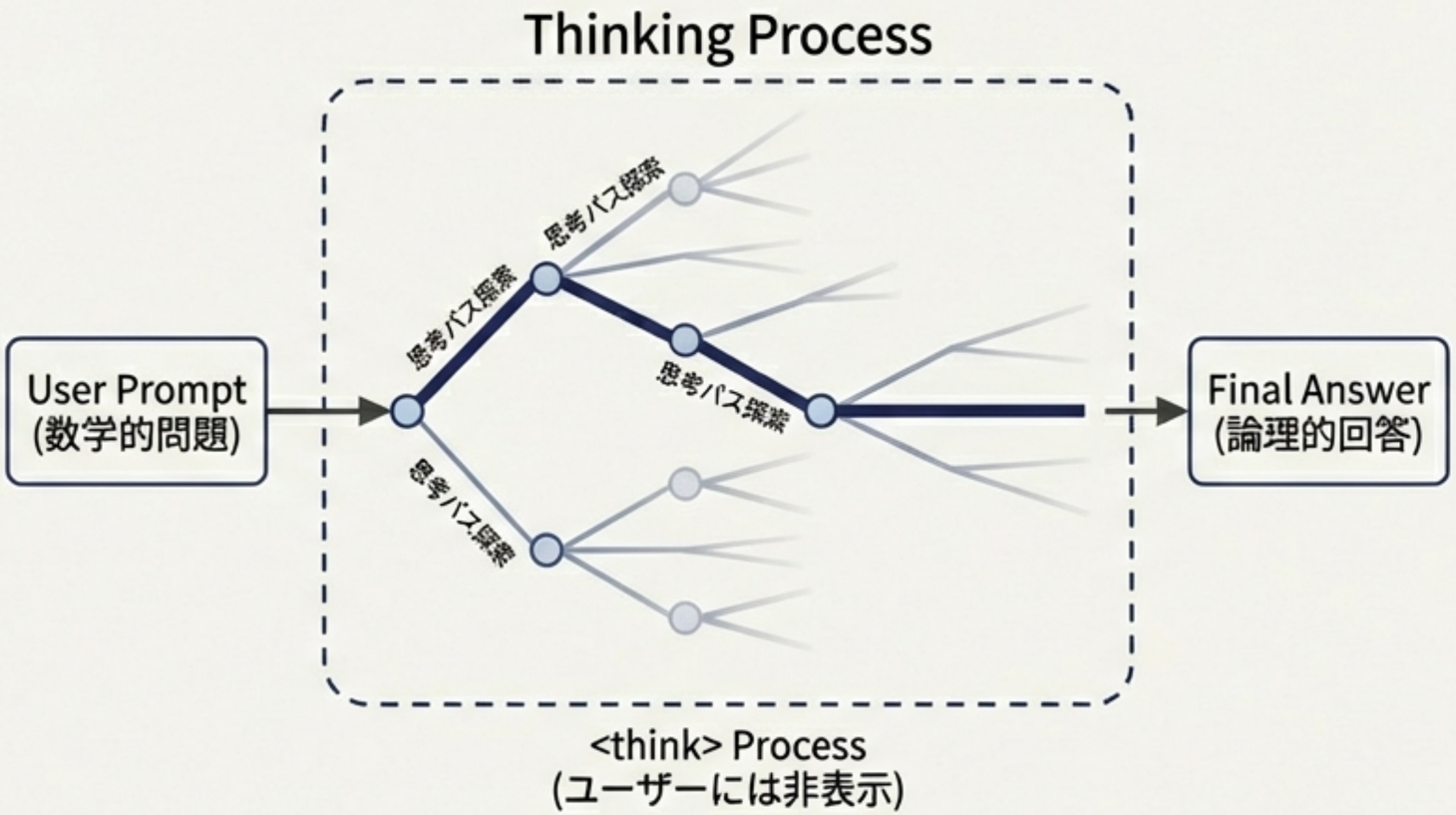
ELYZA：汎用推論モデルへの直接的挑戦

「Thinking」プロセスの実装による論理能力の飛躍

Model: ELYZA-Thinking-1.0
(Qwenベース)

技術的ブレイクスルー:

モンテカルロ木探索（MCTS）と蒸留: データ不足の課題を解決するため、MCTSを用いて「正解に至る思考のパス」を大量に合成。
日本語での推論データを自律的に生成・学習。

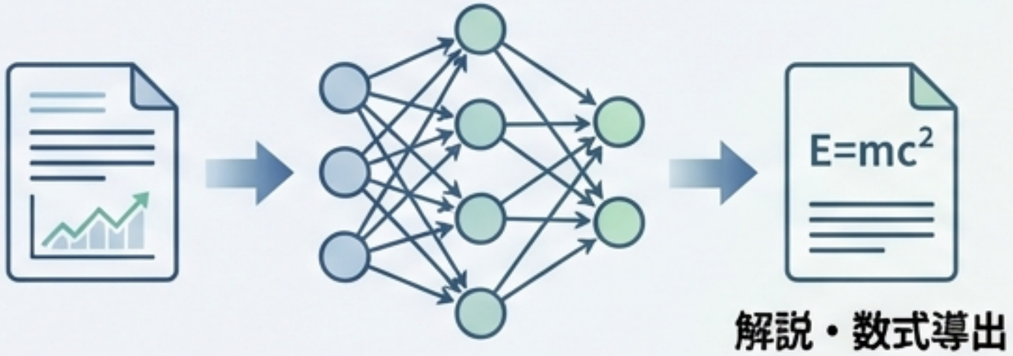


成果： 記述式数学問題において、途中式や論理展開の記述能力が大幅向上。
OpenAI o1などの世界的トレンドに最も近い「正統派」のアプローチ。

Preferred Networks：物理世界の実装

「言語による解説」から「物理現象のシミュレーション」へ

ELYZA / 一般的なLLM (Language)



物理法則を「言語・論理」として処理し、
解説や数式導出を行う。

PFN / Matlantis (Simulation)

物理法則（ポテンシャルエネルギー曲面）を内包し、原子の挙動そのものを計算・予測する。

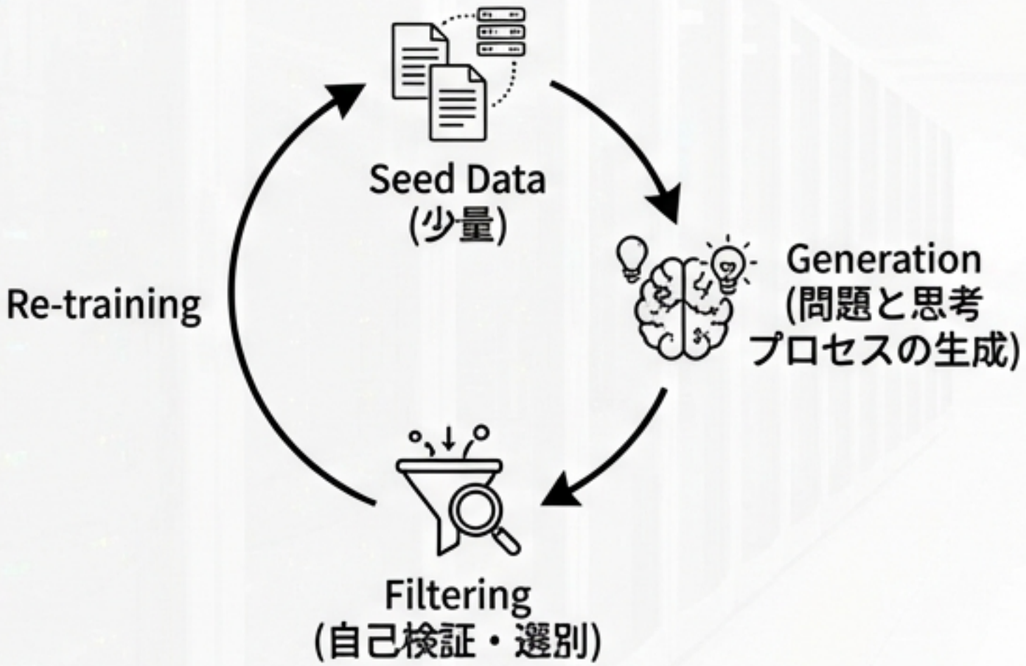
従来の第一原理計算(DFT)比
最大2,000万倍
の高速化

言語モデル「PLaMo」が司令塔となり、Matlantisをツールとして操作する「自律的科学発見」を目指す。

SoftBank：インフラと合成データによる「基礎体力」強化

国内最大級の計算基盤と自己増殖的な学習サイクル

Core Tech: CoT-Self-Instruct



課題: 日本語の推論データ不足。
解決策: モデル自身に思考プロセスを生成させ、厳格なフィル-タリングを経て再学習させるサイクル。

Real World App: Large Telecom Model（LTM）

通信ネットワークという巨大な物理システムの最適化に特化。電波伝搬（物理）とトラフィック（時系列）の複合要因を推論。

NTT & NEC：産業応用と視覚的推論による信頼性

NTT (tsuzumi) : 視覚的機械読解 (Visual Machine Reading Comprehension)

テキストだけでなく、図表、グラフ、回路図を「視る」能力。専門書や論文の図表を理解し、テキスト情報と統合して推論する。

NEC (cotomi): 論理と知識の融合

LLMと「論理推論エンジン」のハイブリッド構成。プラント制御などのミッションクリティカルな領域で、物理法則やマニュアルと矛盾しない「ハルシネーションのない推論」を実現。

Takeaway: アカデミックなスコアよりも、社会実装における堅牢性を重視。

特定領域への特化と、自律型科学者の誕生

CyberAgent: セキュリティ推論（CTI）

Cyber Threat Intelligence（CTI）分野に注力。
断片的なログデータから攻撃者の意図や脆弱性を論理的に「推論」する能力。
数学とは異なるが、高度な論理結合力が求められる領域。

Sakana AI (The Horizon): The AI Scientist

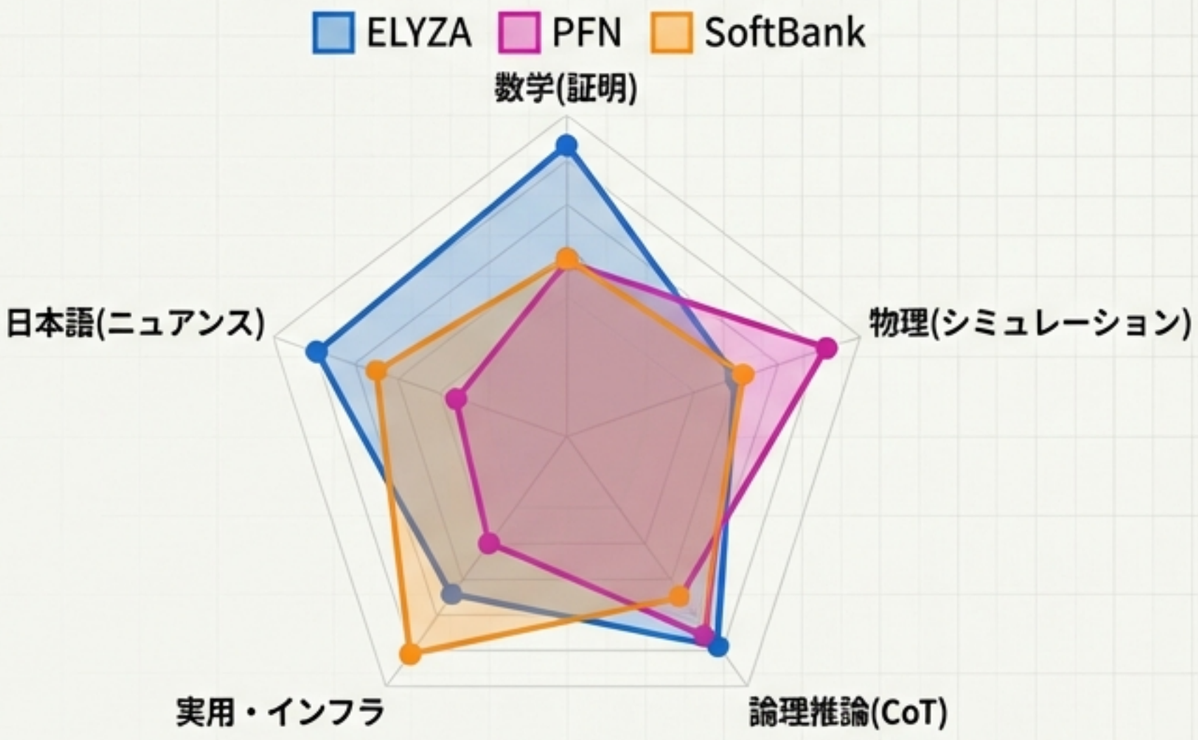
「問題を解くAI」から「研究するAI」へ。
アイデア生成、実験コード作成、実行、結果分析、論文執筆、査読までを全自動で行うエージェント。
日本のAI開発全体が目指す「自律的科学発見」の先駆的事例。

総合比較：分野別推論能力スコアカード（2026）

企業・モデル別推論能力と特性

企業/モデル	数学（証明）	物理（シミュレーション）	論理推論（CoT）	実用・インフラ
ELYZA	◎（最高水準）	○（記述）	◎（特化）	○
PFN	○	●（圧倒的）	○	◎（Agent）
SoftBank	○（向上中）	○（通信物理）	○	◎（基盤）
NTT	△	○（図表理解）	○	◎（視覚）
NEC	△	○（制御）	○	◎（信頼性）

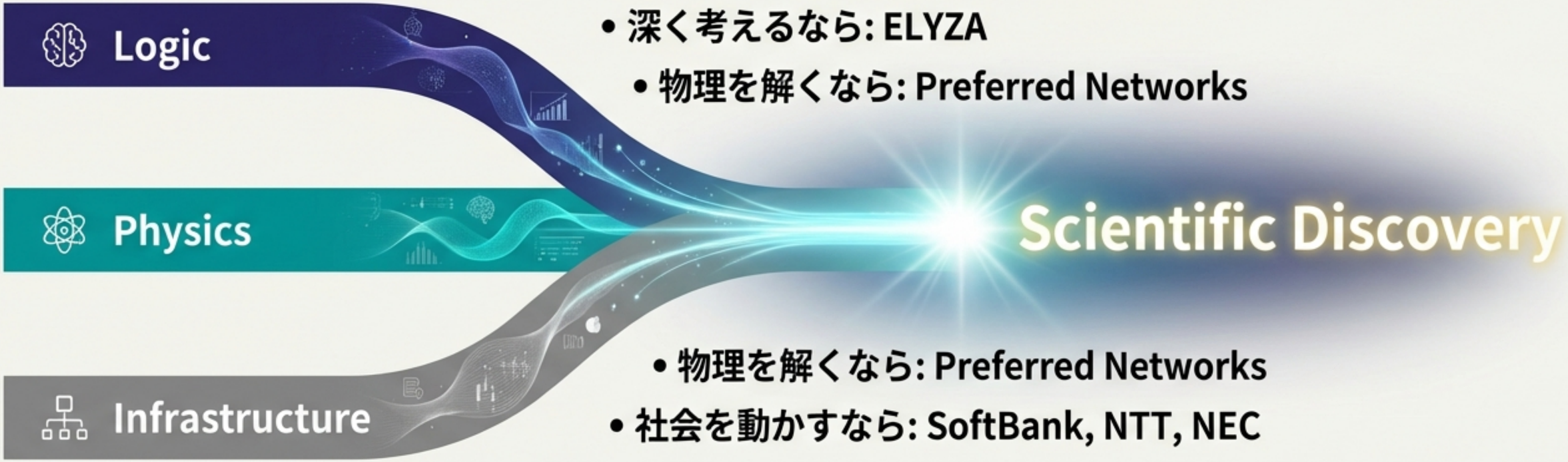
特性比較レーダーチャート



ELYZAは数学的論理に、PFNは物理世界の再現に、大手通信系は産業応用と信頼性に強みを持つ。

結論：日本のAIエコシステムは「科学的発見」へ向かう

競争から共創へ —— 多様な推論能力の統合が鍵



- 深く考えるなら: ELYZA
- 物理を解くなら: Preferred Networks
- 物理を解くなら: Preferred Networks
- 社会を動かすなら: SoftBank, NTT, NEC

これらの多様なモデルが連携（エコシステム化）することで、日本は「AIによる科学的発見（AI for Science）」の領域で世界をリードするポテンシャルを持つ。2026年は、AIが単なるツールを超え、知的な探求者としての地位を確立する元年となる。