ELYZA-Thinking-1.0：国産オープン推論モデルの実力と立ち位置

ELYZAが公開した商用利用可能な国産推論モデル「ELYZA-Thinking-1.0」。NII (国立情報学研究所) の「LLM-jp-4」をベースに開発され、OpenAIのo1-miniを超える推論性能を示し、高い透明性とセキュリティが求められるビジネス利用の新たな選択肢として登場した。

推論水準は「o1-miniを上回り、o1には届かない」位置付け



o1-mini
GPQA 60.00 | AIME 63.60

ELYZA 33Bモデル
GPQA 69.95 | AIME 68.85
主要英語推論ベンチマークで o1-miniを凌駕する。

o1 (正式版)
GPQA 75.7〜78.0 | AIME 76.80

コード生成とツール呼び出しでベースモデルから大幅進化



ベース (LLM-jp-4) 比で LiveCodeBenchが 24.6pt向上。

ビジネス・実務における導入価値



商用利用可能な「Apache 2.0」ライセンスで公開
自社専用環境へ構築でき、高いセキュリティ要件に対応可能。



1〜2人日の人手作業で開発パイプラインを実現
合成データと強化学習により最小限の人手で高性能化を達成。

Gemini Notebook