

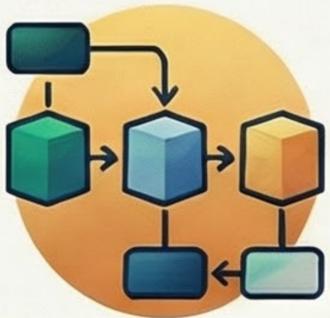
国産初の推論型AI「PLaMo 3.0 Prime」：日本発の“長考”できる思考エンジン

技術的革新：国産初の「推論 (Reasoning)」エンジン



国内初の「System 2」的思考プロセスを実装

単なる次単語予測ではなく、課題を細分化・推敲して解く「長考」が可能。



3段階の高度な事後学習パイプライン

SFT、DPO、強化学習 (RL) の全工程で「思考過程」を学習し、論理性を強化。



長大な文脈を処理する「YaRN」技術の採用

入力最大64K、出力最大20Kトークンの広大なコンテキスト窓を実現。



PLaMo 3.0 Prime

性能比較と実用的な導入メリット



日本語の指示追従・対話性能でグローバルモデルを凌駕

JFBenchやJapanese MT-Benchで、gpt-ossやQwen3を上回る圧倒的性能を記録。

| | PLaMo 3.0 Prime (β版) | gpt-oss-120b | Qwen3-235B (Thinking) |
|-------------------|----------------------|--------------|-----------------------|
| 日本語指示追従 (JFBench) | 0.371 (最上位) | 0.323 | 0.301 |
| 英語指示追従 (IFBench) | 0.663 (最上位) | 0.623 | 0.520 |
| 日本語対話 (MT-Bench) | 9.01 (最高水準) | 8.95 | 8.99 |

主要なベンチマークにおけるグローバルモデルとの性能比較 (PFN公開データより抜粋)



徹底したデータ主権と柔軟な提供形態

セキュアなVPC内デプロイやオンプレミス稼働に対応し、機密情報の漏洩を防ぐ。



2026年6月中旬、商用版リリース予定
現在β版モニターを募集中で、実務トラフィックでの検証と最適化を進行中。