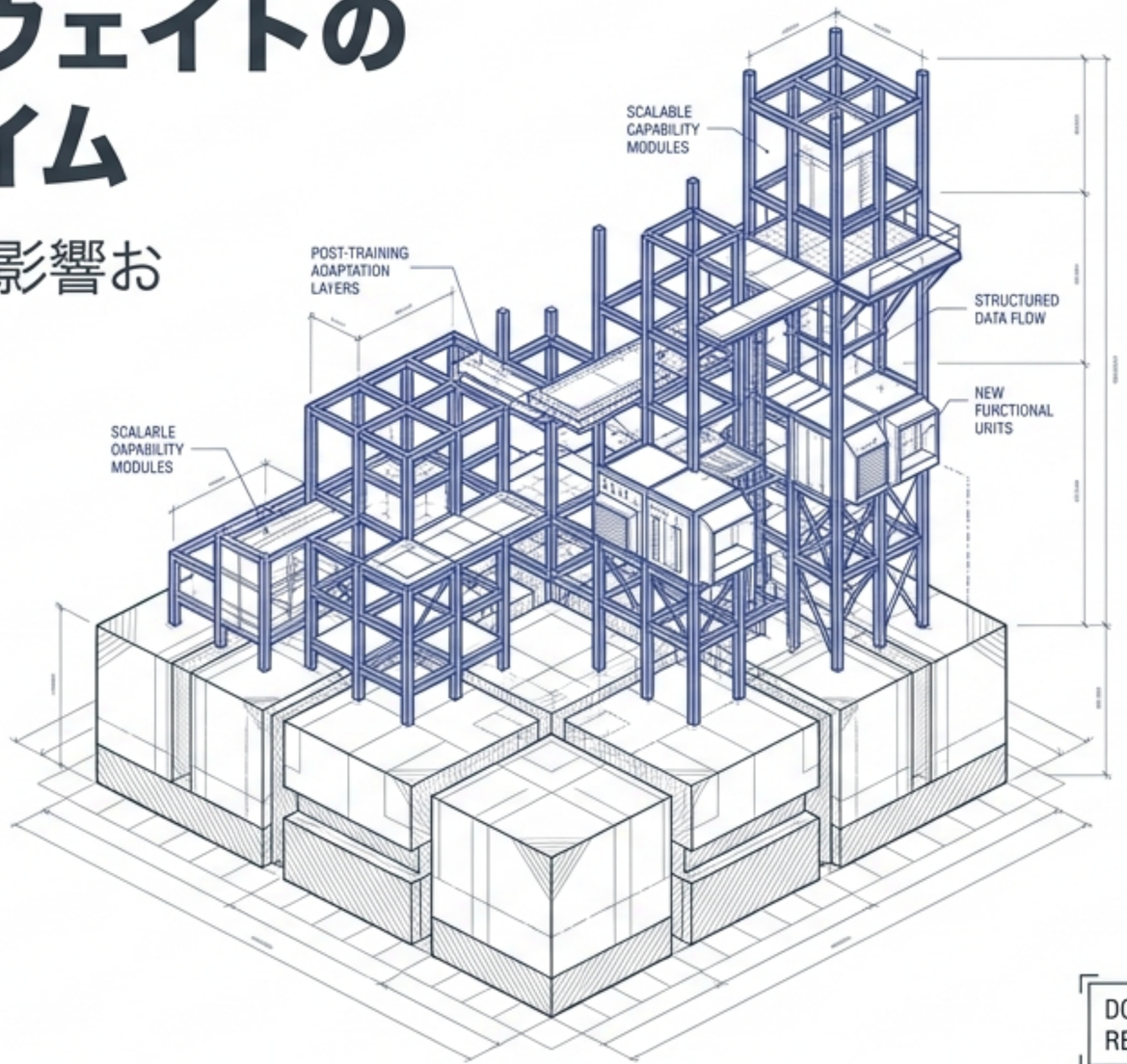
事後学習のスケールアップが切り拓くオープンウェイトの新たなパラダイム

Z.ai GLM-5.3 産業的影響および技術評価レポート



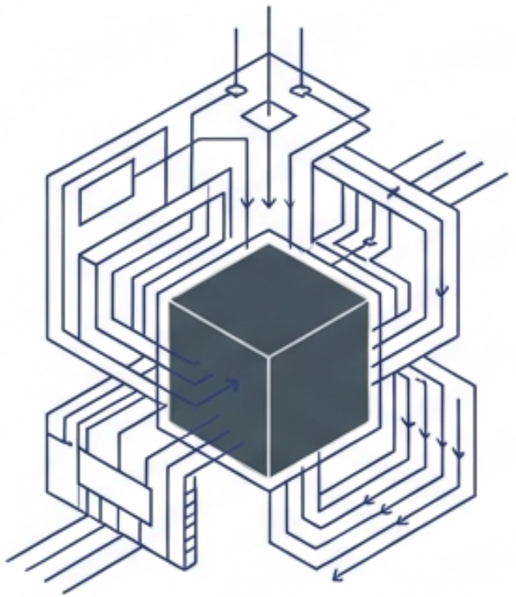
DOCUMENT CLASSIFICATION: INDUSTRY IMPACT REPORT // TARGET: AI CAPABILITY & CYBERSECURITY

パラメータ拡張への依存からの脱却

凍結された基盤モデル

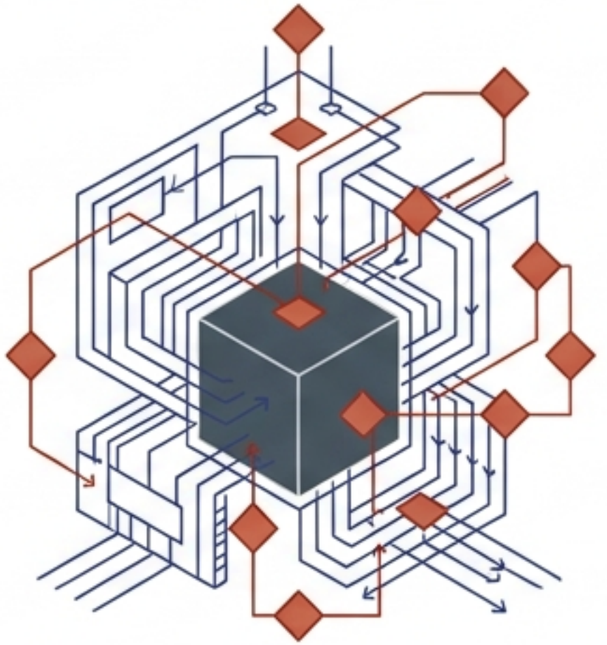
GLM-5.2 アーキテクチャ

7,430億（743B）パラメータを完全維持。ベースモデルの再学習やアーキテクチャ変更は一切なし。



事後学習（Post-Training）の極大化

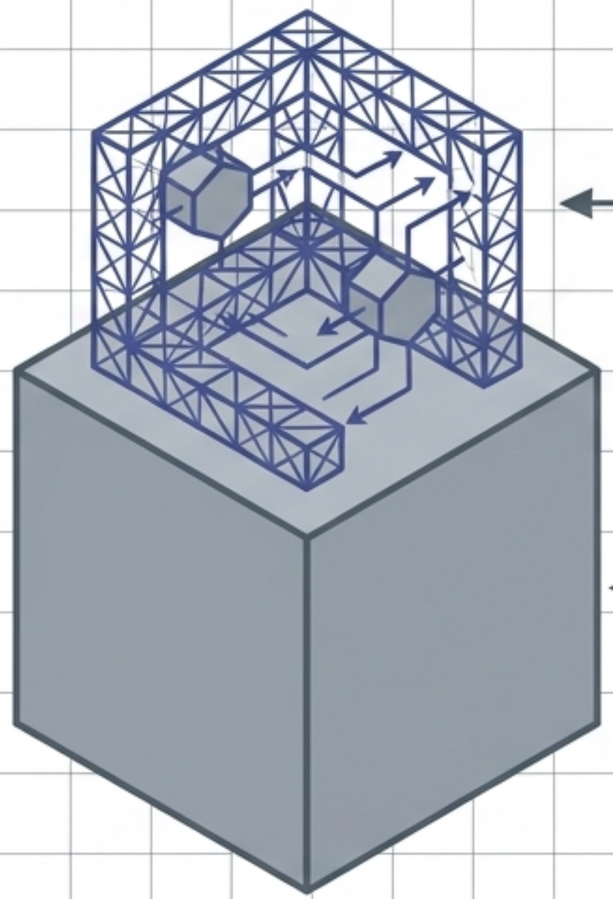
強化学習（RL）と合成環境のスケールアップのみで、ソフトウェア開発能力を前世代比50%向上。



自律的なサイバーセキュリティ能力の伸長

クローズドなフロンティア級セキュリティモデル（Anthropic Mythos 5）を凌駕する脆弱性特定能力の創発。

GLM-5.3 システムアーキテクチャと要求仕様



GLM-5.3
Post-Training
(RL + Synthetic Data)

GLM-5.2
(743B Base Model)

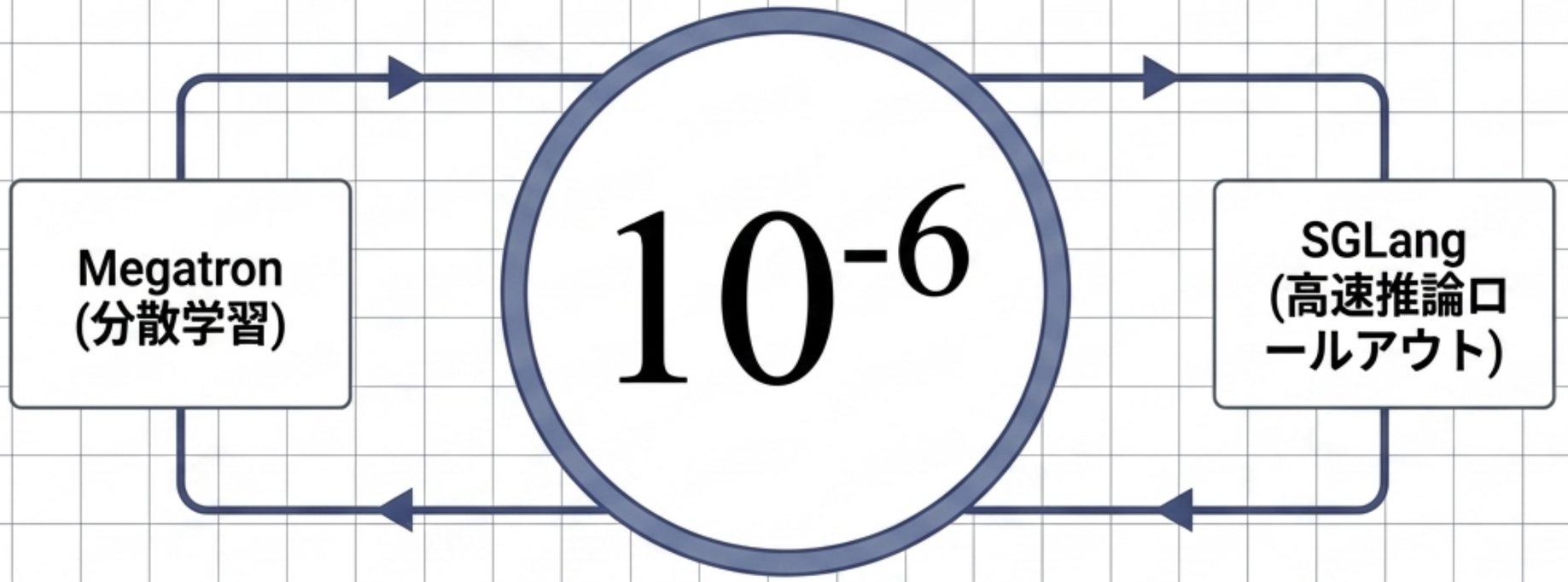
Context Window:
1,000,000 トークン (1M)

Output Capacity:
最大 131,072 トークン (131K)

Reasoning Effort Requirement:
思考機能の完全無効化（disabled）は非対応。API移行時は「low / high / max」のいずれかの思考強度の指定が必須。

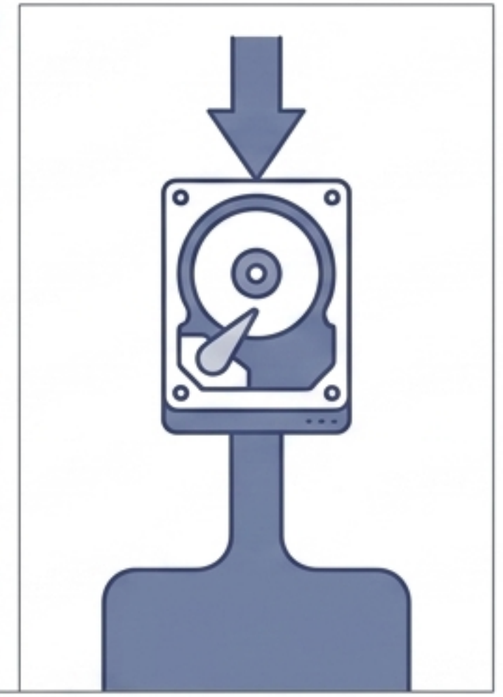
Insight: 既存の長文脈処理基盤「IndexShare」と大規模非同期訓練基盤「slime」を継承。
ハードウェア拡張ではなく、学習インフラストラクチャの最適化が性能飛躍の鍵となる。

訓練パイプラインの収束安定性とコスト削減



対数確率（logprob）誤差の抑制レベル

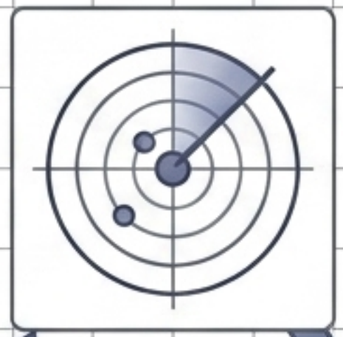
訓練パイプラインと推論ロールアウト間の誤差を極限まで抑え込み、数値的な完全一致を達成。実験の制御精度とRLの収束安定性を劇的に向上。



マルチティーチャーOPD（オンライン・ポリシー蒸留）

ローカルストレージをキャッシュ層として利用する独自手法。長時間タスクにおけるロールアウト計算コストを大幅に低減。

実務特化型タスク環境の自律的生成ループ



Step 1: リサーチエージェント (Research Agent)

人間の実際の開発パターンを分析し、マルチステップ処理や隠れた状態を含む複雑な作業環境を自動生成。

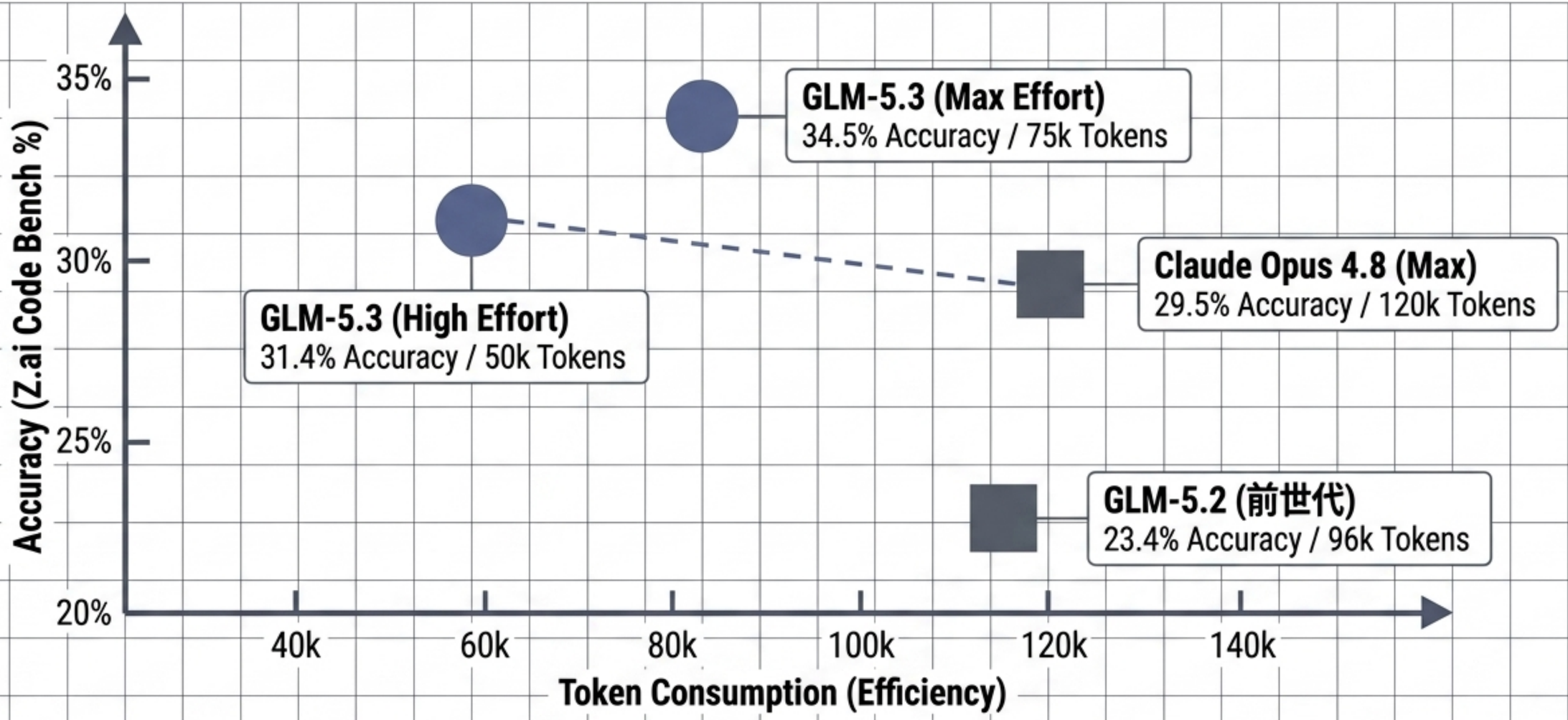
Step 2: ジャッジエージェント (Judge Agent)

生成されたタスク環境の解決可能性を検証し、不正なショートカット（報酬ハック）の有無を即座に判定。

Step 3: 報酬モデル (Reward Model)

厳格な基準をクリアした環境のみを強化学習の報酬として適用し、ベンチマーク暗記からの脱却を実現。

フロンティア級の精度と圧倒的なトークン消費効率



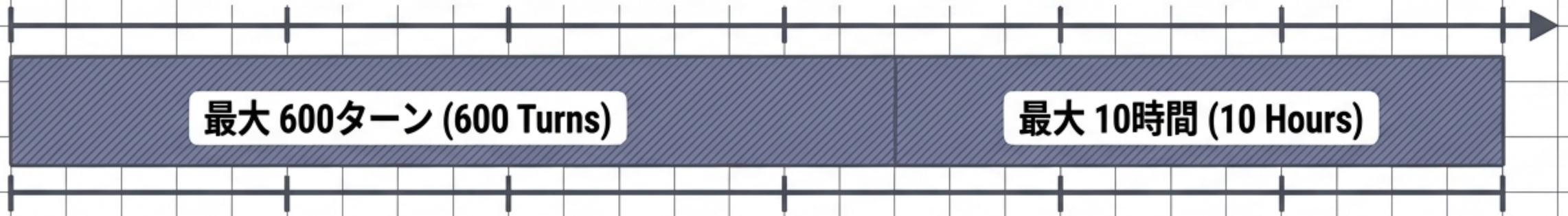
Key Takeaway: High設定のGLM-5.3は、Claude Opus 4.8（Max設定）に対し、精度で上回りながらトークン消費量を半分以下 (120k → 50k) に抑制。

長時間自律エージェントとしての持続的解決能力

Terminal-Bench 3.0 スコア: 28.3

(参考: Claude Fable 5 は 33.74)

Test Environment: Claude Code 2.1.207 を実行基盤として直接運用。



このスコアは単発の対話性能ではなく、長大なコードベースに対する長時間の自律的な問題解決能力を証明している。DeepSWE v1.1においてもスコア66.9を記録し、実務レベルの自律性を確立。

自律発生的サイバーセキュリティ能力の二面性

受動的特定 (CyberGym - 脆弱性監査)	能動的攻撃 (ExploitBench - エクスプロイト推論)
GLM-5.3: 84.5% (Highlight: 制限付きモデルを凌駕)	Anthropic Mythos 5: 78.0%
Anthropic Mythos 5: 83.8%	GLM-5.3: 54.4% (Highlight: 前世代から倍増するも未到達)
GLM-5.2: 77.2%	GLM-5.2: 24.4%

Insight Note: 防御的なコードレビュー能力の強化が先行。特定した脆弱性を実際に動作する攻撃用コードへ変換するには、さらなるRLステップが必要。

リアルワールドでの実証：オープンソースインフラの脆弱性監査

269

対象オープンソースプロジェクト（OSカーネル、ブラウザ、ネットワーク機器等）

2,436 件

特定されたセキュリティ脆弱性の総数

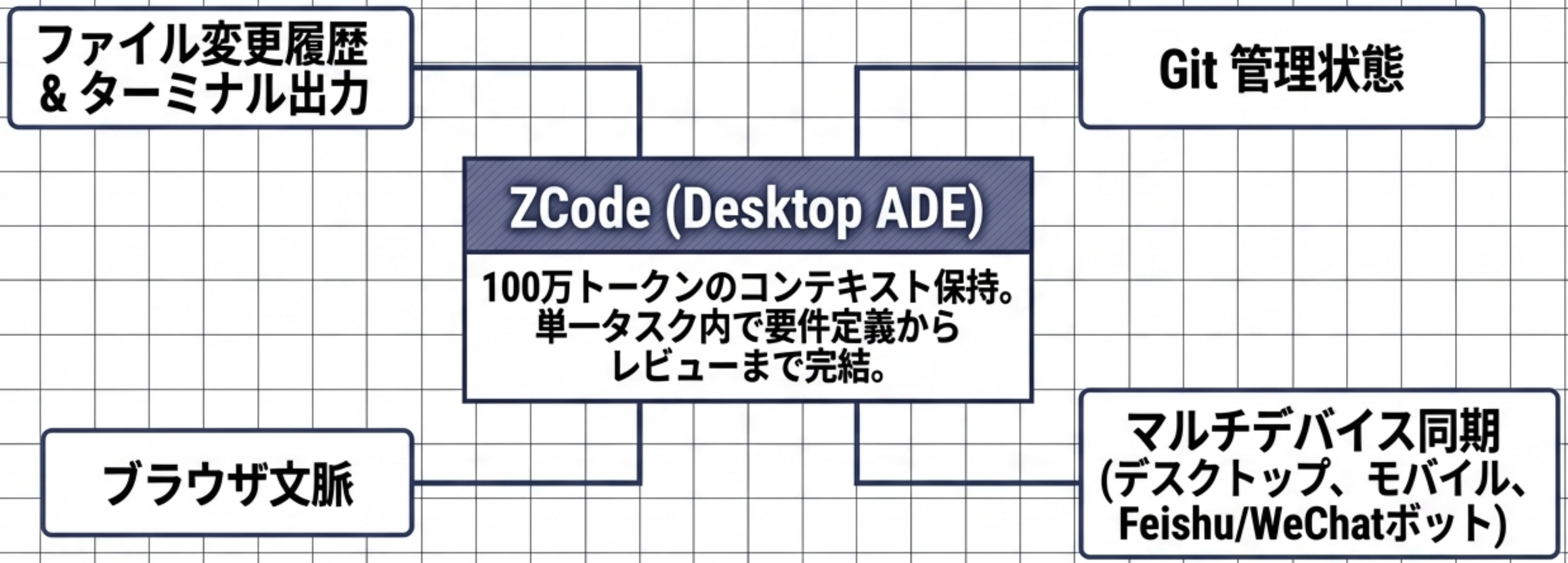
1,097 件

「Critical（緊急）」または「High（高危険度）」の深刻度

平均 26.6 年間

検出された最も古い脆弱性は1981年に作成されたコード内に存在。平均26.6年間見過ごされてきたゼロデイ欠陥をAIが発掘。

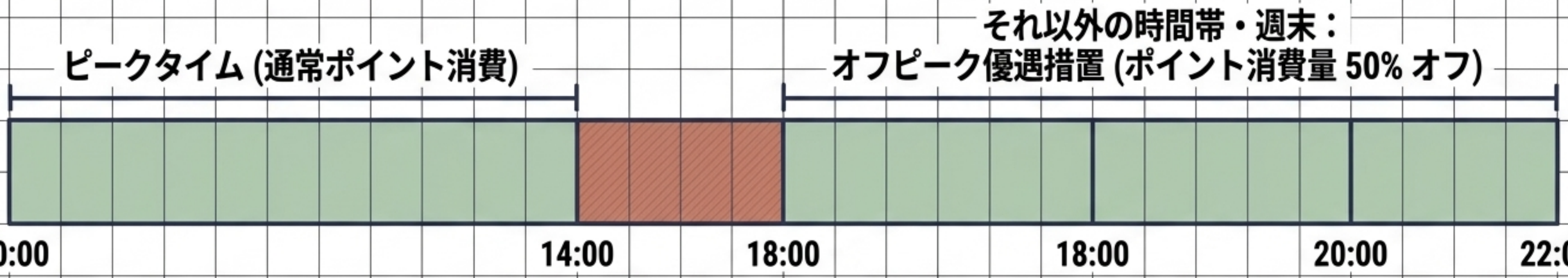
ZCode：深層結合された統合開発エコシステム



Performance Metric: プロンプトキャッシュヒット率 94%〜98% 以上。反復的なコンテキスト読み込みにかかる計算コストとレイテンシを劇的に抑制。

GLM Coding Plan：利用経済性とルーティング戦略

24-Hour Timeline Flow (UTC+8 /JST)



待機時間タスク (Idle Tasks)

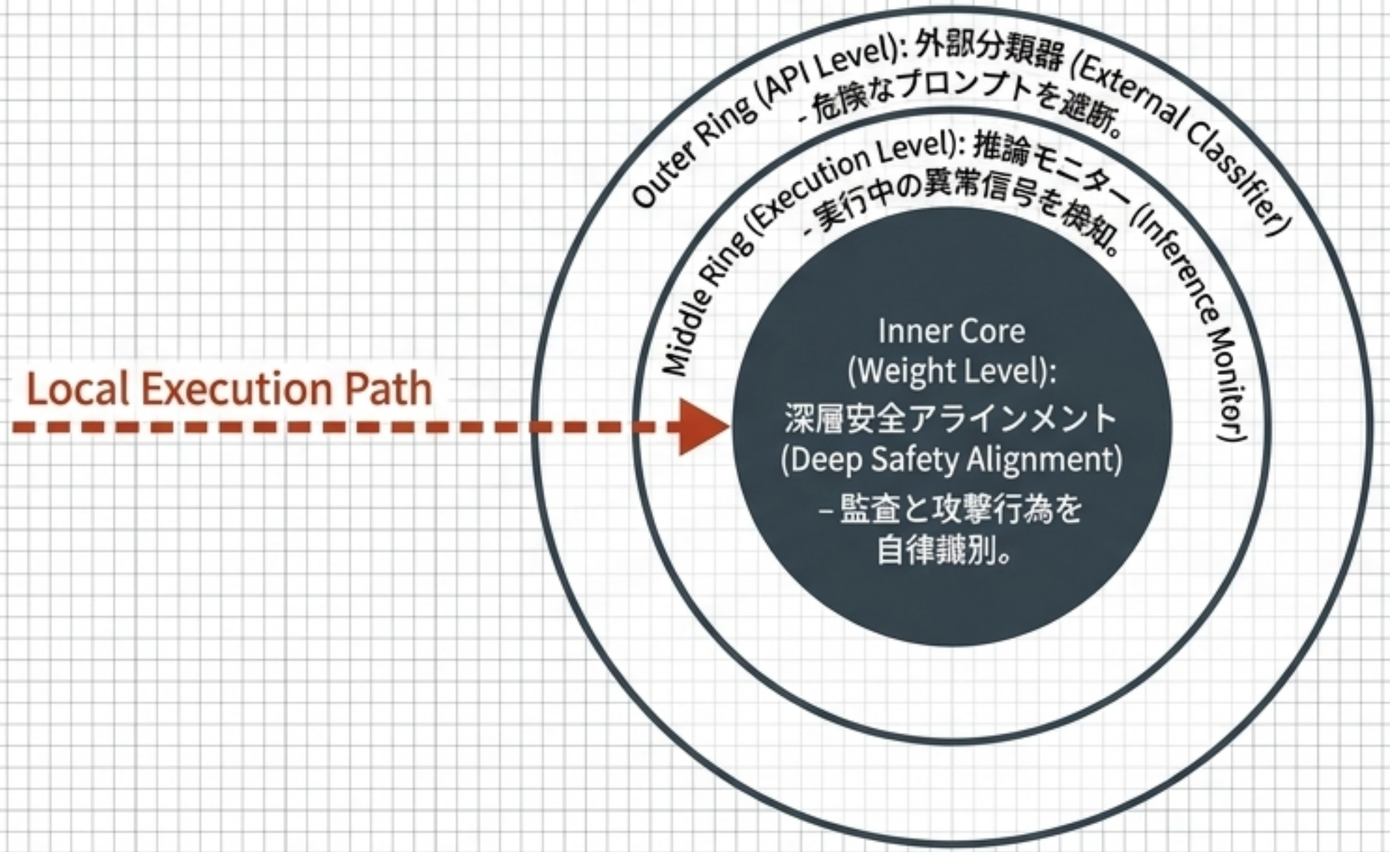
緊急性を要さないバックグラウンド作業をキューに投入。システムの空きリソース発生時に、利用枠を消費せず無料で処理を実行。

Compatibility Note: Anthropicのカスタムベース**URL**設定を利用し、既存のClaude CodeやOpenCodeから直接呼び出し可能。

フロンティア級モデルの普及形態をめぐるパラダイム対立

オープンモデル陣営 (Z.ai)	閉鎖型制限モデル陣営 (Anthropic)
Philosophy	Philosophy
• 防御技術の独占防止と民主化	• 厳格な審査と悪用防止の徹底
Initiative	Initiative
• Open Source Shield (OpenVuln) イニシアチブ	• Project Glasswing (Mythos 5)
Execution	Execution
• 脆弱性の無償監査。発見された脆弱性は公開台帳に記録され、責任ある開示（Responsible Disclosure）を経て修正。	• 攻撃的アラインメントを除去し、審査を通過した特定企業や公的機関にのみアクセスを限定。

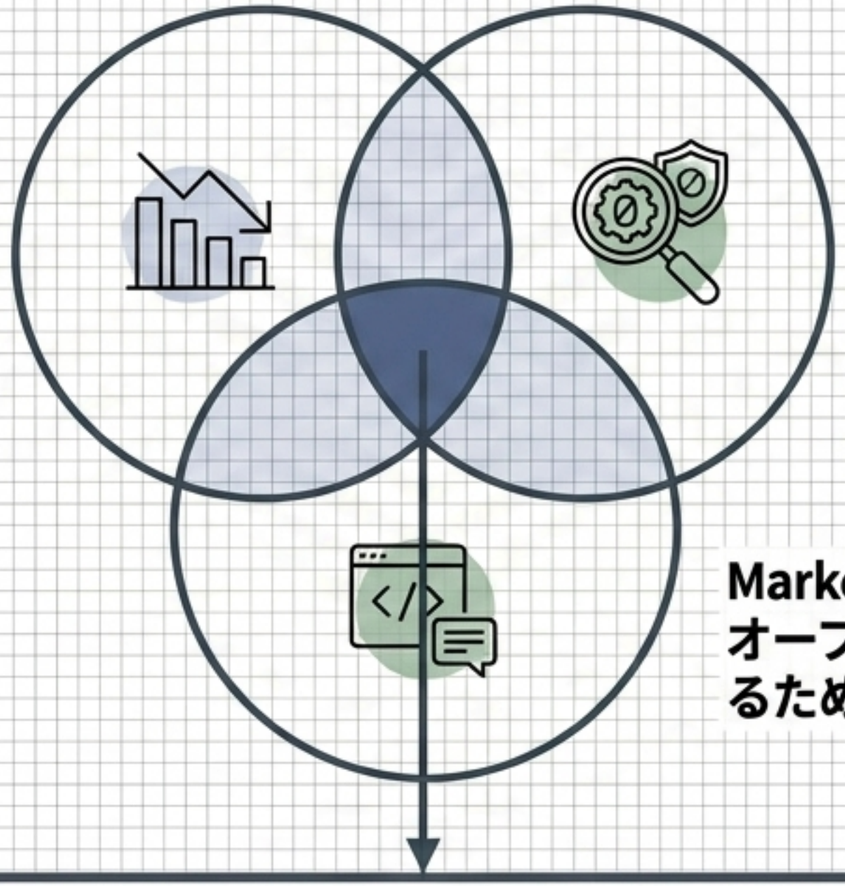
3層セーフガード構造とローカル実行の脆弱性



The Vulnerability: ローカル環境での運用時、クラウドインフラに依存する「外部分類器」と「推論モニター」は機能喪失する。重みが開放されれば、ファインチューニングによるアラインメントの迂回（デュアルユース・リスク）が不可避となる。

デュアルユースのジレンマと新たなROIがもたらす構造的帰結

Technical Reality:
事後学習によるコスト崩壊
（ベースモデル凍結による
開発効率の劇的向上）

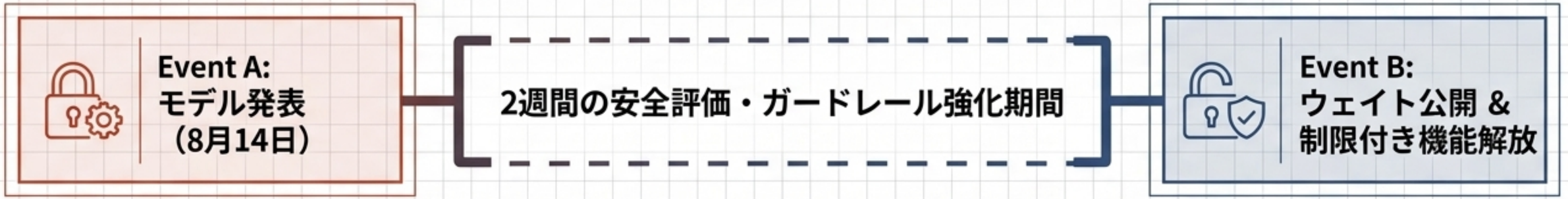


Emergent Capability:
未知のゼロデイ欠陥を自律発見できるサイバーセキュリティ能力

Market Pressure:
オープンソースインフラを防御するための能力の民主化要請

フロンティア性能への到達コストが激減したことで、最高峰のセキュリティ能力を民主化することが「経済的に可能」となった。しかし、その民主化自体が、オープンウェイトモデルにおける「責任ある公開（Responsible Release）」の全く新しい基準を業界に強要している。

中国AIガバナンスの成熟と「信頼済みアクセス」の夜明け



Precedent

中国のAIラボが「安全性とリスク管理」を明示的な理由として、オープンウェイトの公開を段階人化・遅延させた初の実例。

Trusted Access

最も機微な能動的攻撃機能は「信頼済みアクセス（Trusted Access）」承認ユーザーにのみ限定提供。

Future Outlook

この段階的公開モデルが、今後のオープンソースAIコミュニティにおける「責任ある公表」の世界標準となるかが試されている。