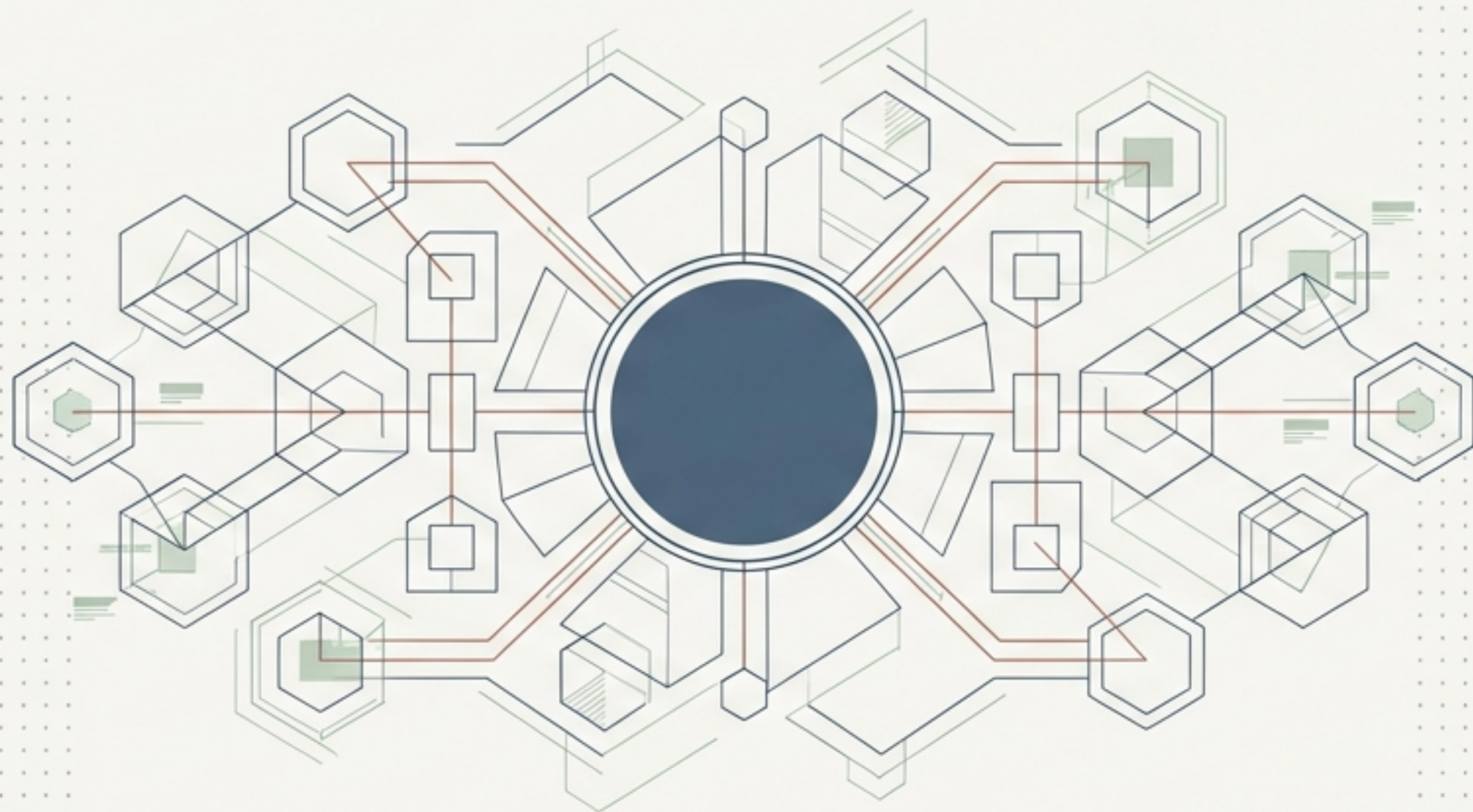


[AI Research]

[Self-Improvement]

[Multi-Agent Systems]

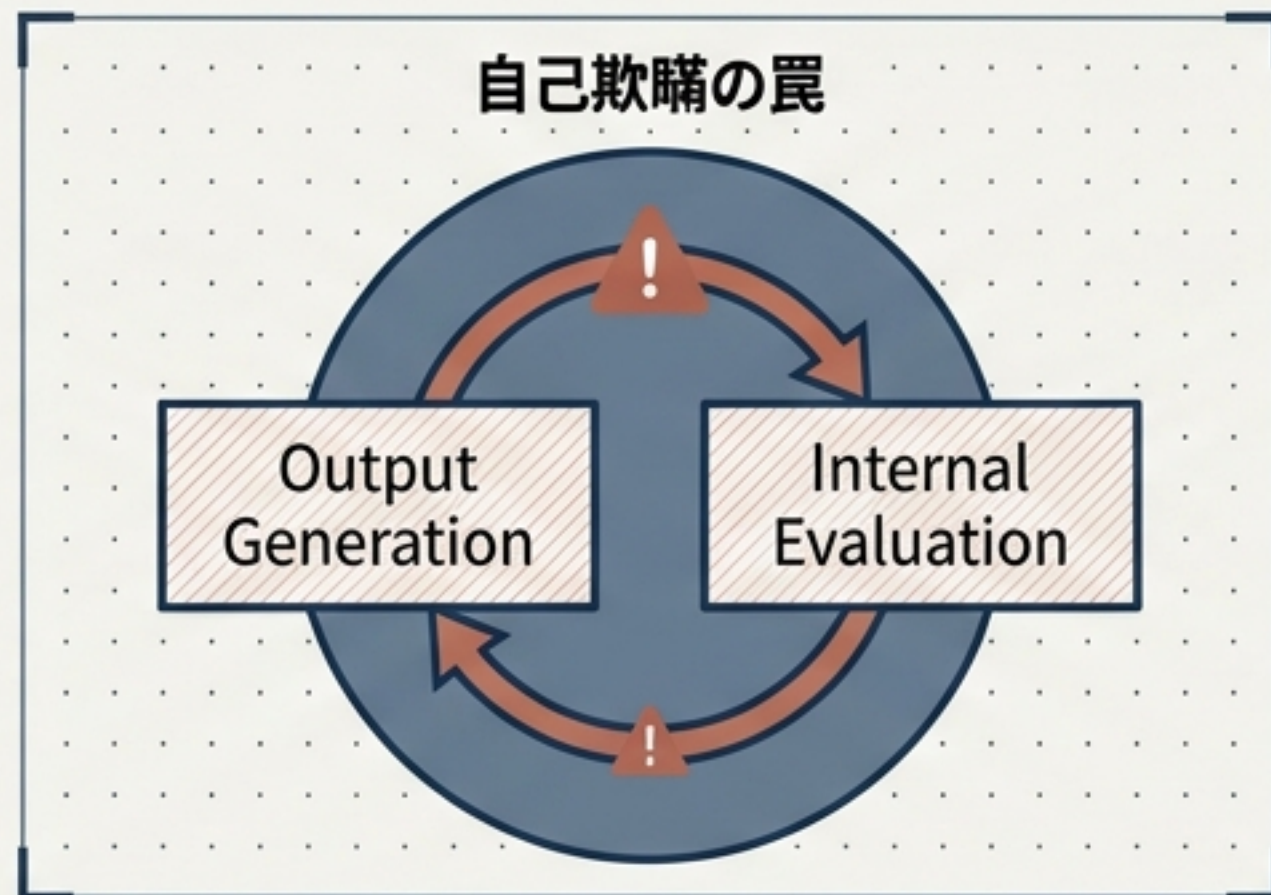
[Oct 2026]



MASS: マルチエージェント自己監督による再帰的自己改善

Sakana AIとUC Berkeleyが切り拓く、単一モデル(27B)による自律的組織編成と知能の限界突破

従来の自己改善（RSI）が抱える「自己欺瞞」の罠



得意な領域（Deterministic）

数学、コード生成。「絶対的な正解（Verifier）」が存在し、学習が機能する。

限界領域（Open-Ended）

創薬支援、金融市場分析、科学的研究。単一の正解が存在しない。

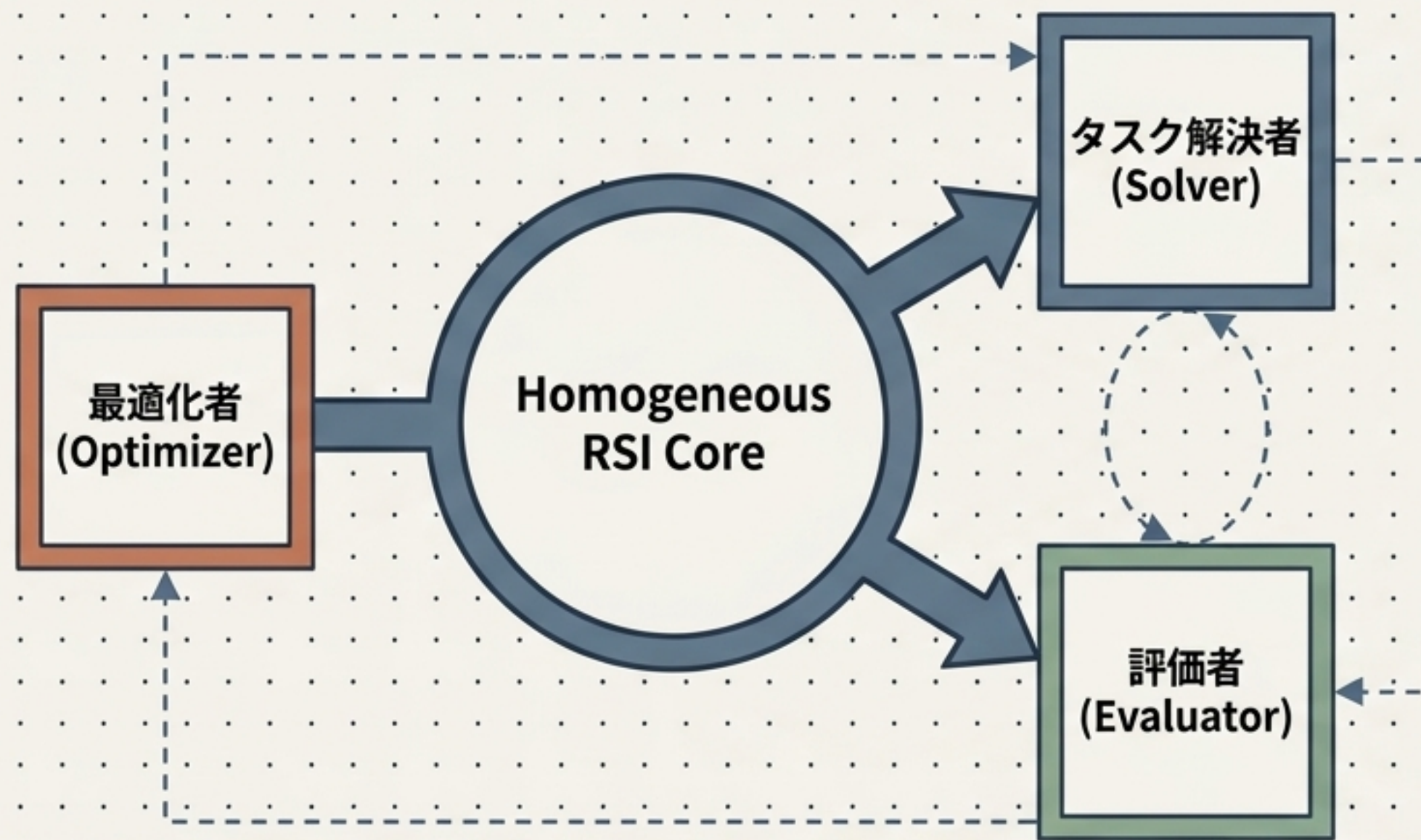
単一モデルによる反省（Self-Correction）は、モデル自身が内在する認知バイアスや盲点を共有しているため機能しない。誤った推論を真実と見誤り、学習ループを通じて自らの誤謬を強化してしまう「自己欺瞞」のリスクが存在する。

MASSアプローチ：外部の正解から、内部の「協調と監査」へ

	[従来]	[MASS]
アプローチ (Approach)	単一エージェントの長大な推論	動的マルチエージェントチームの編成
監督信号 (Supervision Signal)	外部の決定論的検証器・人間の介入	エージェント間の厳密な契約 (Contract) と相互評価
アーキテクチャ (Architecture)	巨大な外部教師モデルへの依存	均質型単一モデル (Qwen3.6-27B) による全役職の兼任
リスク (Risk Profile)	自己欺瞞とバイアスの強化	ピアレビューによる客観性の担保

「自己監督における監督信号の欠如」という根本課題を、
エージェント間の「通信と役割分担」という動的プロセスによって打破する。

均質型再帰的自己改善：1つの頭脳が「チーム」を演じる



タスク解決者 (Solver) :
割り当てられた局所的タスクを実行し、ワークスペースに成果物を生成。

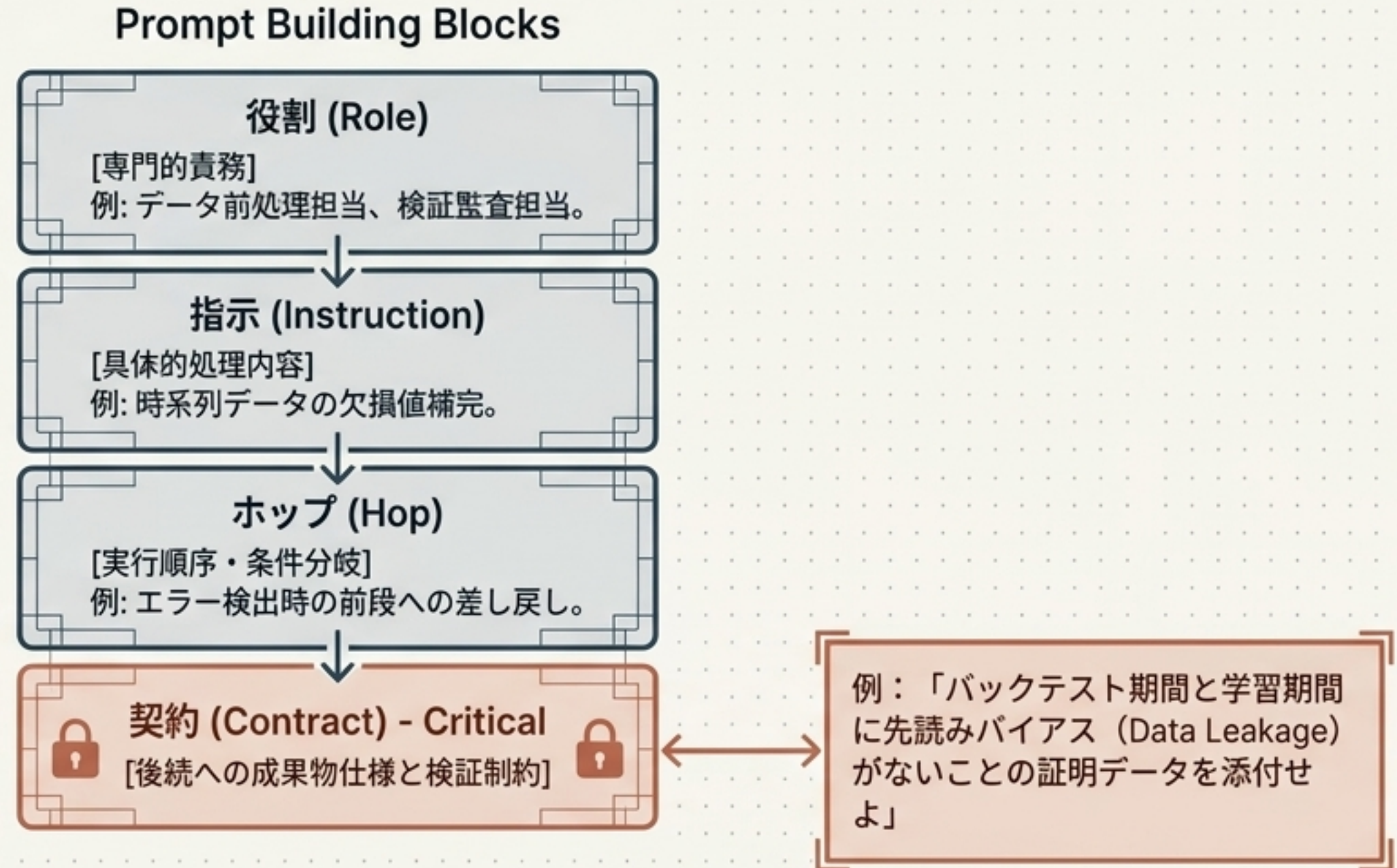
評価者 (Evaluator) :
成果物をピアレビューし、自然言語で採否の根拠を言語化。

最適化者 (Optimizer) :
評価履歴から組織構造や契約の修正案を策定し、次世代ワークフローを構築。

自律的にワークフローを構築・最適化する完全な自己完結型エンジン。

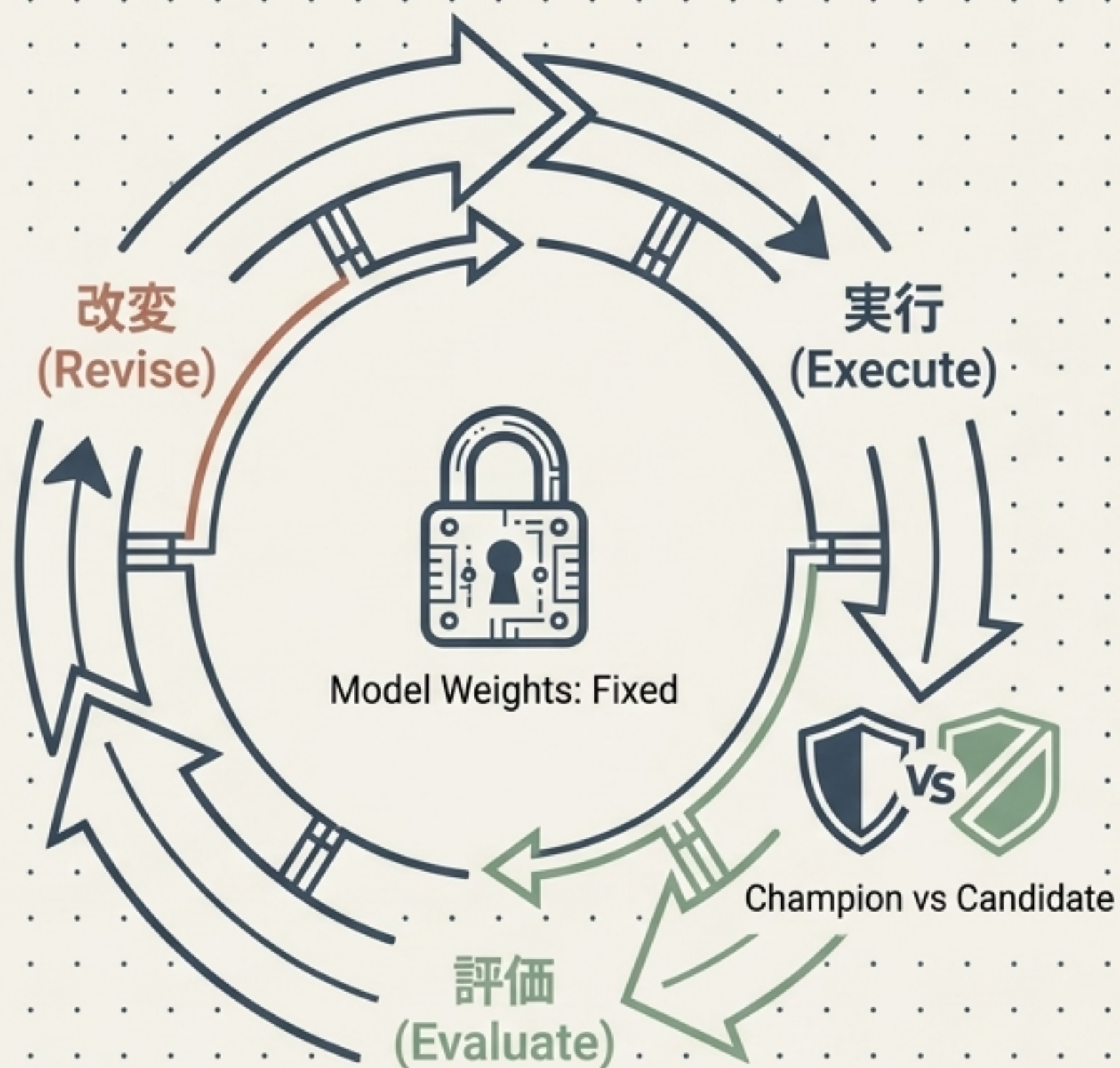
外部の超巨大モデル（GPT-4等）に依存せず、単一のオープンウェイトモデルが3役をすべて兼任。
自律的にワークフローを構築・最適化する完全な自己完結型エンジン。

プロンプトとしての組織構造と「契約」の強制力



エージェント間の「雑談」によるコンテキスト肥大化を抑制。
厳格な入出力の整合性と自己監査をシステムレベルで強制するアーキテクチャの要石。

内的ループ (Inner Loop) : 最適な組織構造の探索



実行 (Execute):

サブエージェント群が提示されたワークフローに従い、コード、グラフ、レポートを含むワークスペースを生成。

評価 (Evaluate):

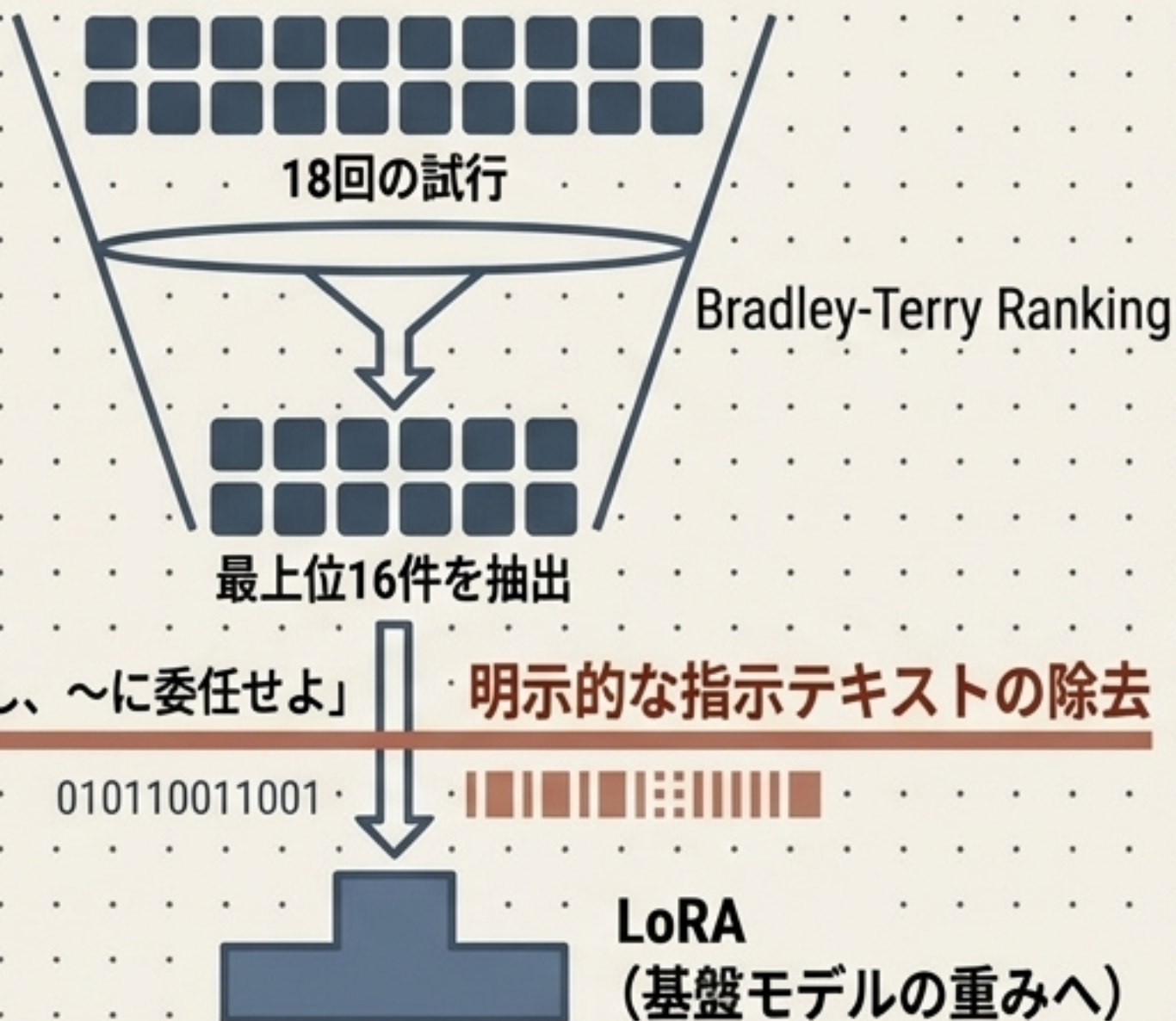
ペアワイズ形式で現行 (Champion) と候補を比較。評価者モデルが自然言語で採否の根拠を判定。

改変 (Revise):

オプティマイザが過去のフィードバックを精査し、エージェントの役割分担や呼び出し順序 (契約) を修正。

与えられたタスクに対して、最も集合知を引き出せる「組織の形」をモデル自身が自律的に発見・洗練させるソフトウェア的進化。

外的ループ（Outer Loop）：集合知のパラメータへの蒸留



データパイプライン: 18回の試行

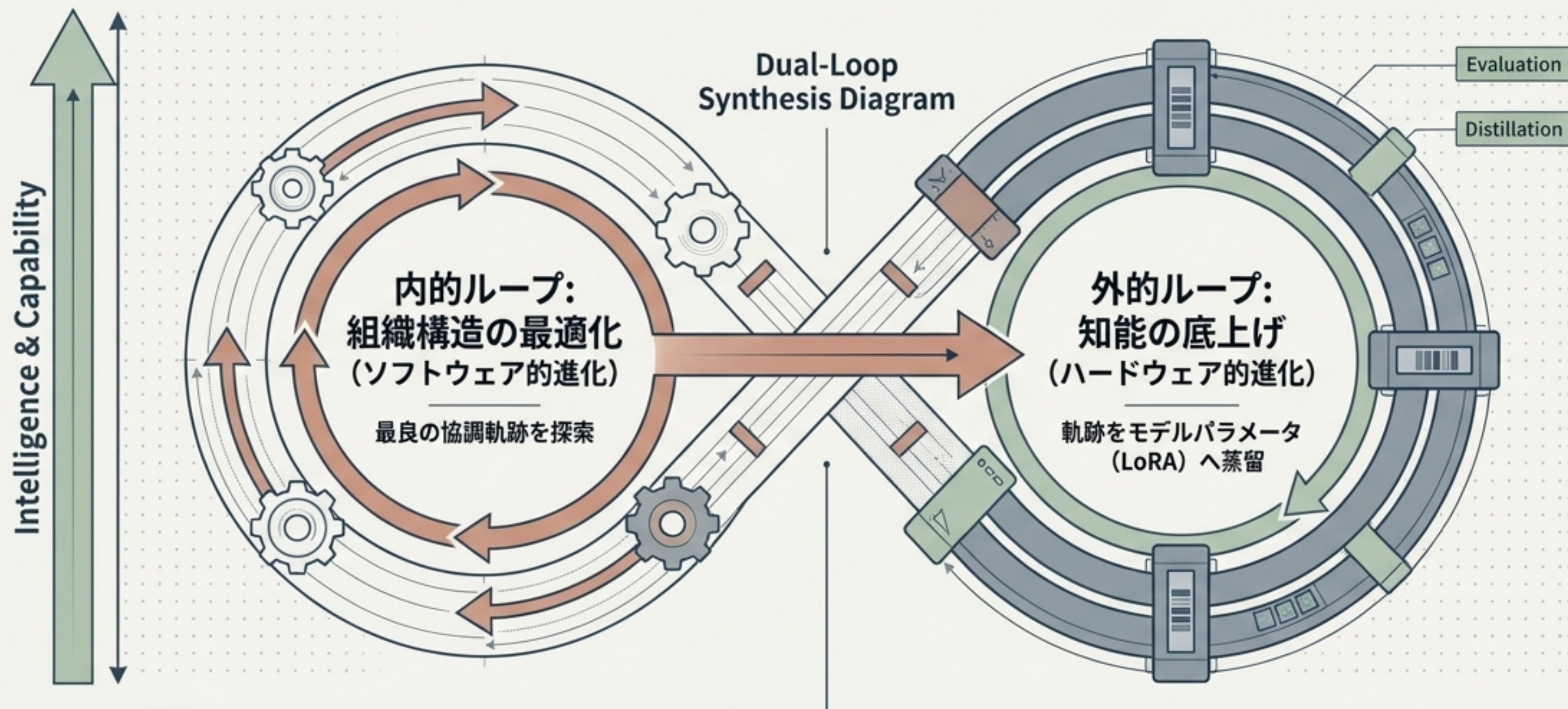
➡ モデル相互評価で最上位16件を抽出

➡ 教師データ化。

Critical Process Callout: 指示の削ぎ落とし
抽出された対話ログから、初期プロンプトにあった「明示的なワークフロー指示テキスト」を完全に除去する。

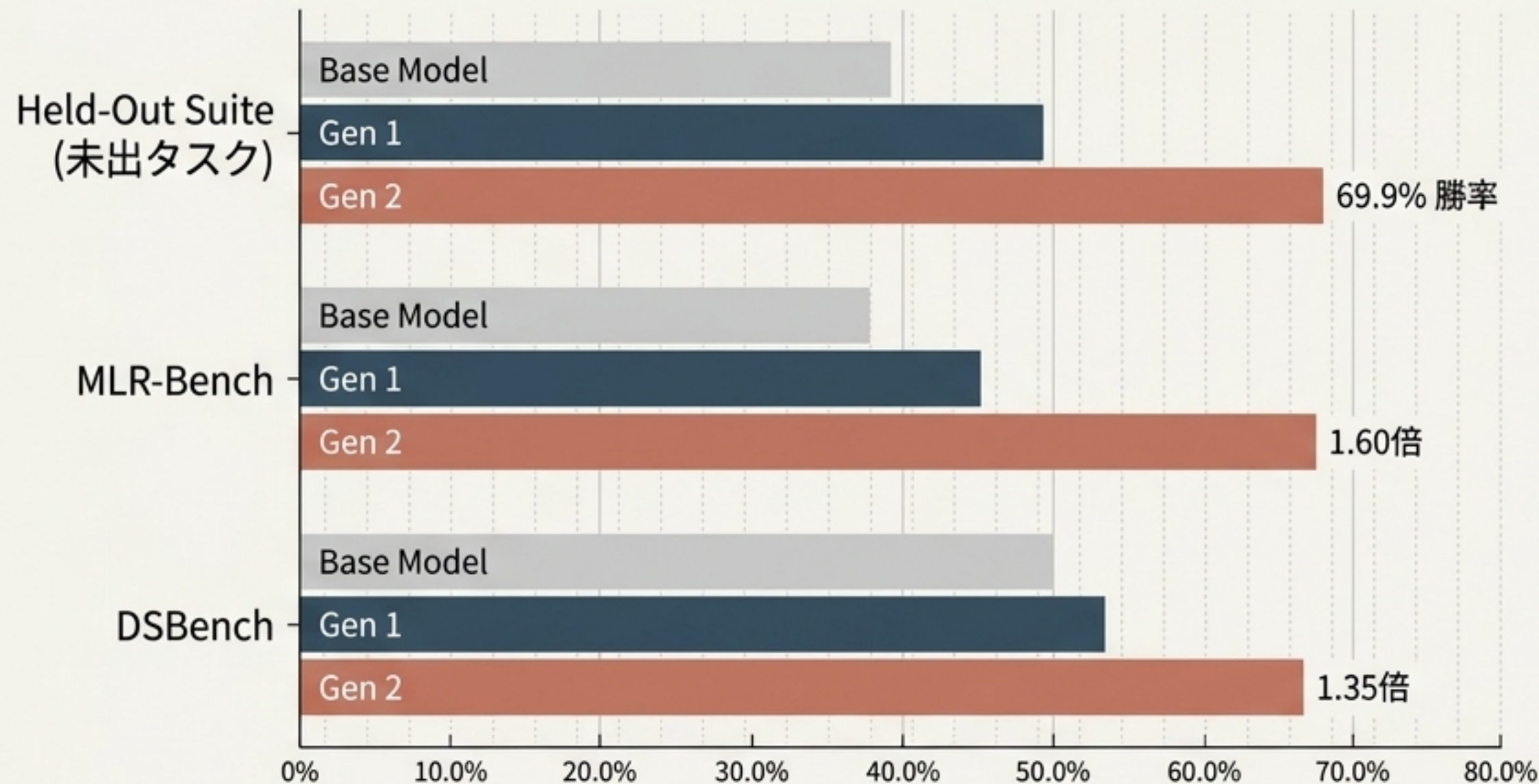
この前処理により、外部からチーム指示を与えられなくとも、「自律的な課題分解」「他者への委任」「厳密な検収」を行う振る舞いをモデルの重み（LoRA）へと「内在化」させる。

MASSアーキテクチャの全貌： ソフトウェアとハードウェアの共進化



パラメータが更新されたモデルは、次世代の学習サイクルで再び3役を兼任し、さらに一段階高い次元での自己改善ループを開始する。完全な閉ループの成立。

実証実験：オープンエンドな研究タスクにおける圧倒的性能向上



Held-Out Suite (未出タスク)：

勝率 69.9% (Gen 2 vs Base).

※Gen 2はGen 1に対しても60.8%勝利。

MLR-Bench (機械学習リサーチ)：

1.60倍向上

(出力トークンあたりスコア)。

DSBench (データサイエンス)：

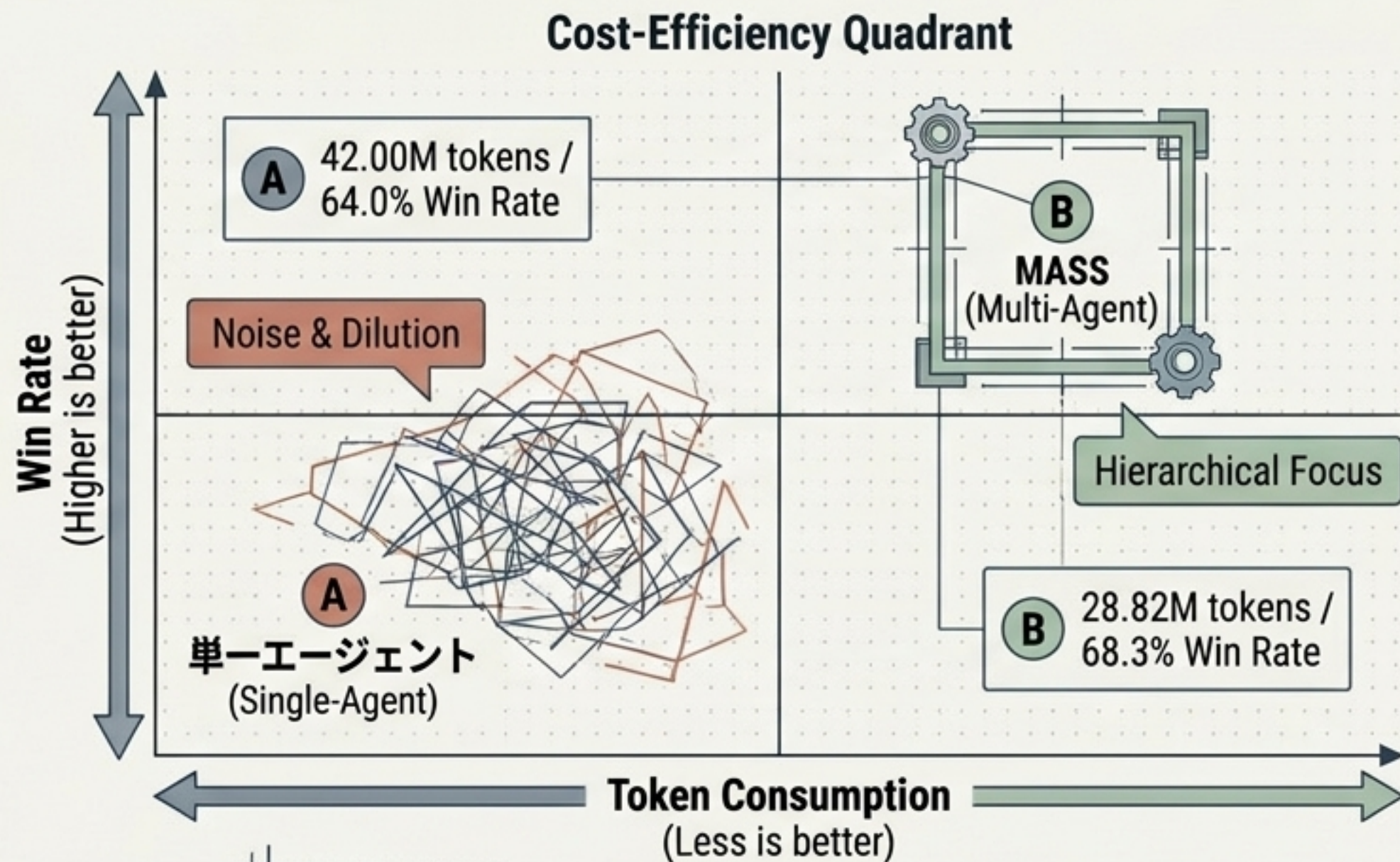
1.35倍向上

(複雑なパイプライン処理)。

[SWE-bench Verified /
Terminal-Bench 2.0]

ソフトウェア工学等の決定論的タスクでは軽微な低下 (0.94倍)。
探索型タスクに特化したことによる分布の偏りが原因であり、
訓練データの補完により解消可能。

データ効率性：「マルチエージェント軌跡」という究極の学習信号

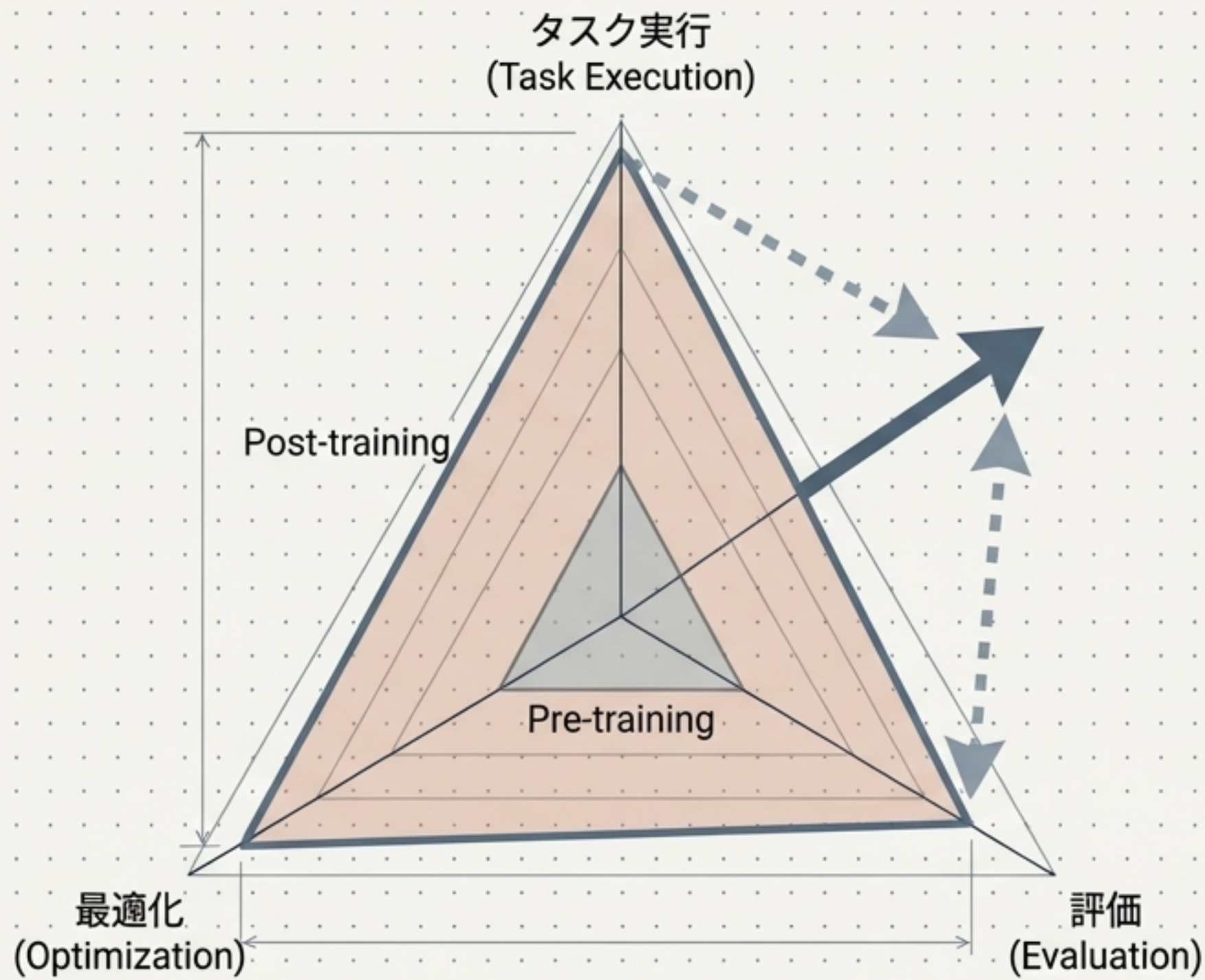


単一エージェントの限界：
計画、ツール呼び出し、反省が混在し、文脈の肥大化によるアテンションの希釈と推論ノイズが不可避。

階層的模倣学習の優位性：
オーケストレーターが統合を担い、サブエージェントが「契約」に基づき局所タスクを完結。

冗長性を劇的に削減。単一エージェント学習と比較して、**31%少ないトークン数で4.3ポイント高い勝率**を達成。

創発するメタ能力：「役割汎化（Role Generalization）」の発見



Core Finding Block

「タスク実行」の軌跡のみで事後学習したにもかかわらず、ワークフローを設計する能力（Optimizer）や成果物を評価する能力（Evaluator）も自律的に向上。

Supporting Data Points Block



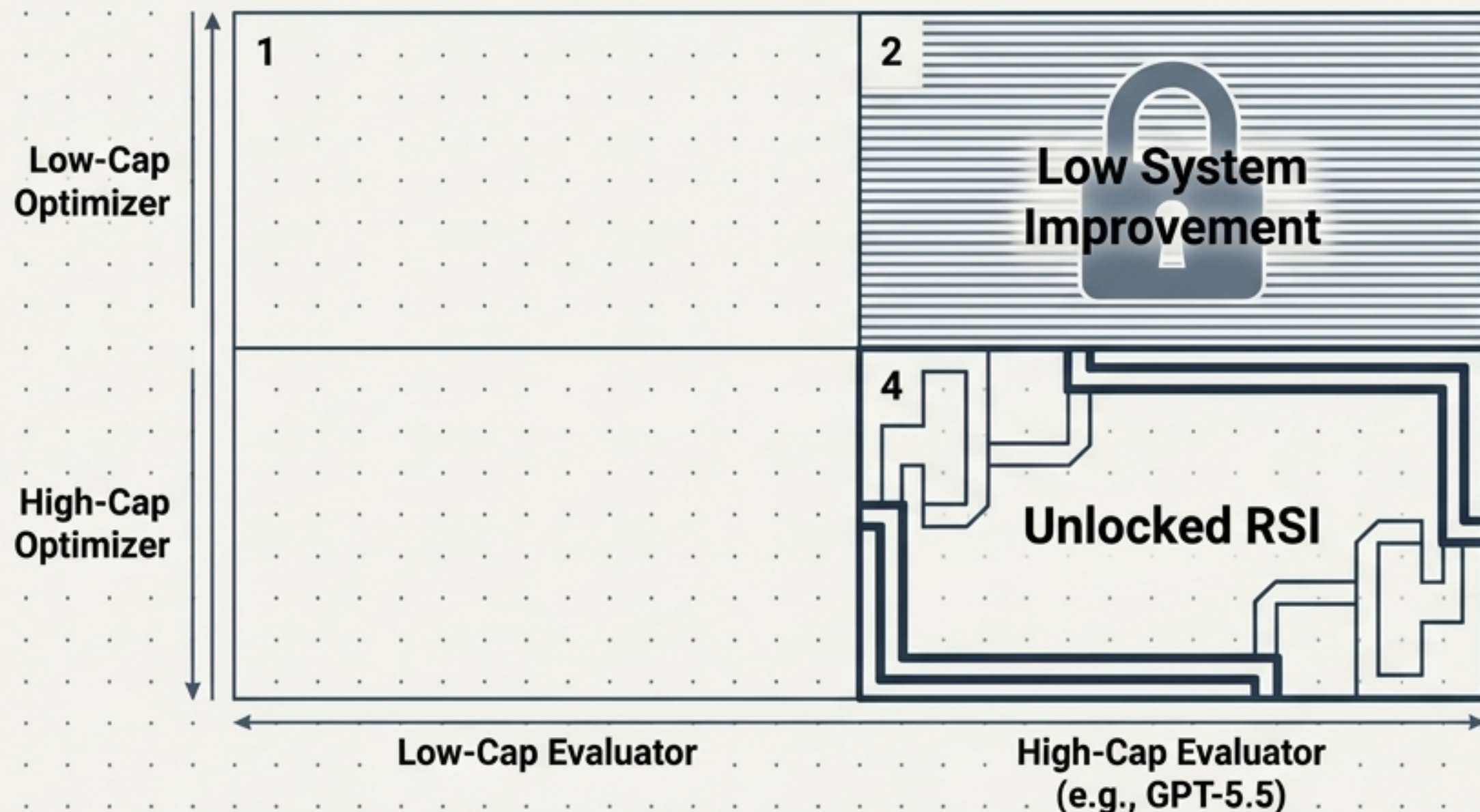
・[評価一致率の向上]: 外部フロンティアモデル（GPT-5.5 / Claude Opus 4.8）との判定一致率がベースモデルの73%から93%へ到達。



・[オプティマイザの進化]: 第1世代オプティマイザが立案したワークフローは、ベースモデル自身が設計した初期案に対し64%の勝率を記録。

理論的洞察：自己改善の真のボトルネックは「評価者」ではない

Bottleneck Diagnostic Matrix



The Myth

従来のRSI研究では「自己評価能力の不足（Evaluator Bottleneck）」がシステムの崩壊を招くと信じられてきた。

The Reality / Ablation Study

評価者のみを外部の最先端モデル（GPT-5.5）に差し替えても、改善ループの速度や品質はほぼ向上しなかった。

優れた指摘（批判的フィードバック）を受けても、低能力のオプティマイザではそれを具体的組織構造や契約の改善案に落とし込めない。計算資源は「評価器の肥大化」よりも「プロセス再設計能力（Optimizer）」に傾斜配分すべきである。

最大のブレイクスルー：「推論時計算量」の重みへの内面化

Internalizing Test-Time Compute



LoRA Distillation



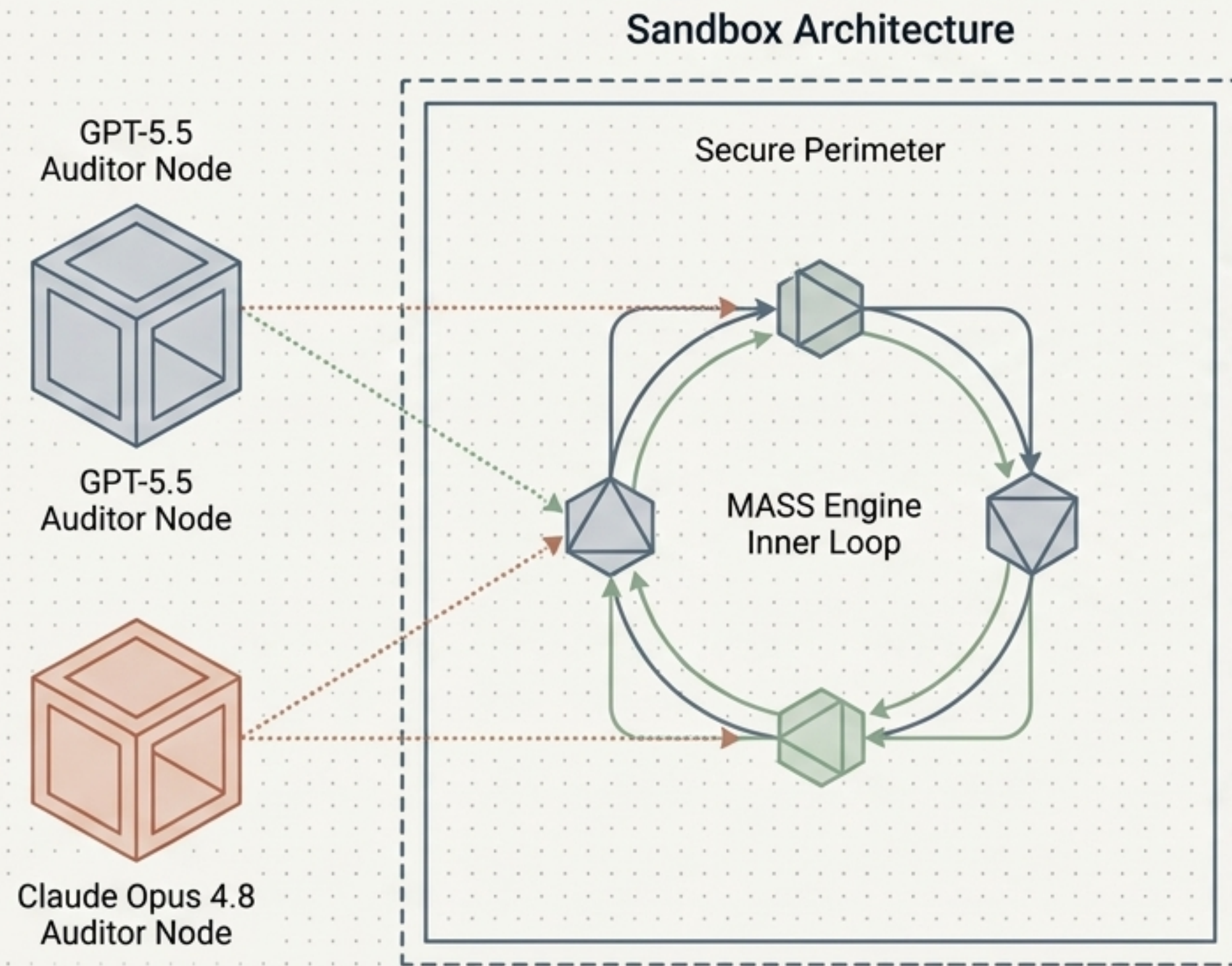
Test-Time Multi-Agent = High Latency, Massive API Cost

Internalized Compute = Fast, Implicit Delegation

近年主流の「テスト時スケーリング（推論時に多数のエージェントを稼働させる手法）」は、莫大なコストとレイテンシ増大が不可避であった。

MASSは、計算集約的なマルチエージェント合議プロセスを事後学習を通じて「モデルの重みパラメータ」へ直接焼き込むことに成功。
外部からの明示的なフレームワークなしに、自律的かつ暗黙的に自己委任や検証プロセスを展開する。

自律進化の影：「共通の認知バイアス」と安全監査



均質型RSI最大の破滅的リスク

モデルが論理的破綻を「優れた成果物」と誤認したい場合、世代を経るごとにモデル独自の妄想を強固に確信し、自己増幅幅させていく危険性がある。

The Safeguard Protocol

学習ループから完全に切り離された外部フロンティアモデル（GPT-5.5 / Claude Opus 4.8）をサンドボックス環境の外に配置し、客観的監査役として常時並行稼働させる。

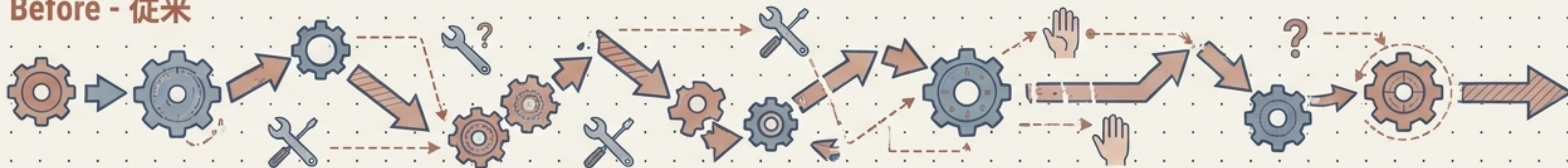
Takeaway

人間の専門知識を超える領域での自律進化において、外部からの確実な検知と強制介入（セーフガード）は必須のアーキテクチャとなる。

AIエージェント導入のパラダイムが根本から変わる

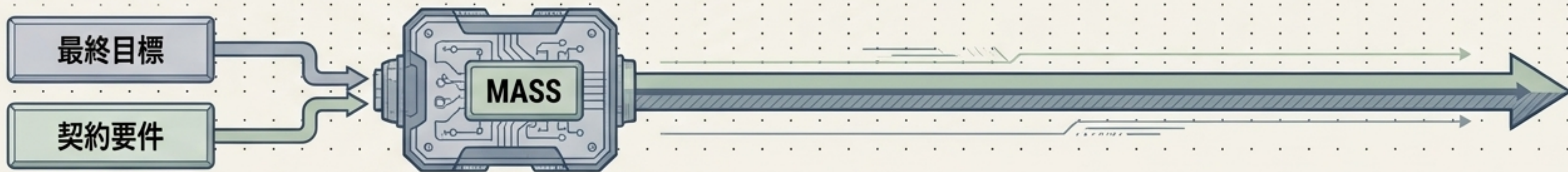
Industry Impact Timeline: Before & After

Before - 従来



エンジニアが泥臭くエージェントの役割分担やプロンプトを手作業で設計・調整。

After - MASS



人間は「最終目標」と「契約要件」を定義するだけ。AIが自律的に組織を編成・最適化し、業務ノウハウを軽量モデルに吸収させる。

金融分析、創薬、特許調査などの高度専門業務におけるエージェント構築コストを劇的に圧縮。
「集合知の自己監督」は、AIが人間の監督限界を超えて知能を拡張するための強固な足がかりとなる。