

# AIエージェントを用いた投資チーム構築：包括的設計・実装レポート

作成日: 2026年2月17日 プロジェクト: finance — 金融市場分析・コンテンツ発信支援ライブラリ 対象読者: プロジェクト関係者、投資運用チーム、AI/MLエンジニア

## エグゼクティブサマリー

### プロジェクトの目的

本プロジェクトは、外国株式アクティブファンドの運用チームにおいて、ファンドマネージャー Y（吉沢）の投資判断ロジックと暗黙知をAIエージェントに体系的に落とし込み、最終的にはマルチエージェント・システム（Multi-Agent System; MAS）による自律的な投資判断の再現・拡張を目指すものである。

### 背景と動機

伝統的なアクティブ運用では、人間のアナリストとファンドマネージャーによる非構造化データの分析が意思決定の中核を担ってきた。しかし、この手法には情報処理能力の物理的限界、損失回避バイアスや確認バイアスといった認知的欠陥、そして暗黙知の伝承困難性という構造的課題が存在する。近年、BlackRock の AlphaAgents、Rochester Institute of Technology の MASFIN、UCLA/MIT の TradingAgents といった学術研究が示すように、LLM（大規模言語モデル）を中核としたマルチエージェント・システムは、これらの課題を構造的に克服する可能性を持つ。

### 現在の到達点

プロジェクトは4段階（Stage 1~4）で計画されており、**Stage 1（Y の判断軸の抽出・文書化）**が完了している。具体的には以下の成果が得られた：

- **Y の却下パターン 7類型の特定**：結果と原因の取り違え、一般論、差別化不明、事実誤認、飛躍的解釈、定量裏付け欠如、戦略と優位性の混同
- **Y の高評価パターン 6類型の特定**：定量的裏付け、直接的CAGR接続メカニズム、能力>結果、明確な因果チェーン、他社事例との類推、構造的vs補完的の区別
- **Y のブランド力評価フレームワーク 4類型の抽出**
- **現行システムの課題 6項目の特定**

並行して、9本の最先端論文の調査・分析が完了し、マルチエージェント・アーキテクチャの設計知見が蓄積されている。

### 学術研究からの主要知見

調査した9本の論文から、以下の設計原則が導出された：

1. 専門エージェント間の議論（Debate）がハルシネーションを抑制する（AlphaAgents, TradingAgents）
2. 5段階パイプライン（Postmortem→Screening→Analysis→Timing→Portfolio）が体系的バイアスを排除する（MASFIN）
3. 数理最適化（HedgeAgents）と強化学習（FinRobot）がポートフォリオ配分を動的に最適化する
4. 先読みバイアス排除とPoint-in-Timeデータ管理がバックテストの信頼性を担保する（StockBench, Glasserman et al.）
5. LLMによるSeed Alpha自動生成が、SSE 50で53.17%の累積リターンを達成した（HKUST）

### 統合設計の方針

本プロジェクトでは、Y の投資哲学（実践知）と学術研究の知見（理論知）を統合した独自のマルチエージェント・アーキテクチャを構築する。核となるのは以下の8エージェント構成である：

FM Agent (Y's Philosophy) — 全体統括・最終判断

— Bull Agent

— Bear Agent

— CA Critic

— CAGR Validator

— Causality Checker

— Historical Analogy

— Bias Detector

— 買いの根拠を徹底探索

— 売りの根拠を徹底探索

— Y の判断軸に基づく競争優位性批判

— 数値の根拠と一貫性検証

— 結果 vs 原因の区別を強制

— 過去判断事例との整合性チェック

— 確認バイアス・初期判断固執を検出

次のステップ

- **短期（3ヶ月）**：Critique Agent プロトタイプの構築と Phase 2 データでの検証
- **中期（6ヶ月）**：弁証法的マルチエージェントの実装と既存 finance パッケージとの統合
- **長期（12ヶ月）**：MASFIN 5段階パイプラインの完全実装とフルポートフォリオ最適化

本レポートの構成

部	タイトル	概要
第1部	プロジェクト背景と目的	起源、Y の判断プロセス、既存システム、資産棚卸し
第2部	FM Y の投資哲学の抽出と体系化	スコアリング分析、却下/高評価パターン、ブランド評価、暗黙知、課題
第3部	学術研究サーベイ	9本の論文の体系的レビュー
第4部	マルチエージェント・システムアーキテクチャ	チーム構成、議論メカニズム、合意形成
第5部	5段階投資判断パイプライン	Postmortem→Screening→Analysis→Timing→Portfolio
第6部	アルファ生成とポートフォリオ最適化	Seed Alpha Factory、動的ウェイト、数理最適化
第7部	バイアス排除とバックテスト設計	先読み/生存者バイアス、PiTデータ、評価指標、チェックリスト
第8部	本プロジェクトへの統合設計	Stage 1~4 の具体的設計、パッケージ統合、HITL
第9部	実装ロードマップ	短期/中期/長期計画、技術スタック
第10部	リスクと課題	技術/運用/規制リスク
付録	参照資料	論文一覧、数式、用語集

## 第1部: プロジェクト背景と目的

### 1.1 プロジェクトの起源

外国株式アクティブファンドにおけるAI活用の背景

本プロジェクト「AI Investment Team」は、外国株式アクティブファンドの運用チームにおいて、ファンドマネージャー Y（吉沢、以下 FM Y）の投資判断ロジックと暗黙知を AI エージェントに実装し、AI による投資判断の再現と拡張を目指す取り組みで

ある。

アクティブファンド運用において、ファンドマネージャーの判断力は最も重要な資産でありながら、最も属人化しやすい要素でもある。優秀なファンドマネージャーが長年の経験で培った「何が真の競争優位性か」「どの分析が投資判断に値するか」といった判断基準は、多くの場合、本人の頭の中にのみ存在する。この暗黙知（Tacit Knowledge）を形式知に変換し、さらに AI エージェントとして実装することで、判断の再現性・拡張性・一貫性を確保することが本プロジェクトの核心的な目的である。

FM Y の投資判断ロジックをAIに落とし込む目的

FM Y は、AI が生成した競争優位性分析レポートに対して、鋭い批評と確信度スコアリングを行ってきた実績がある。しかし、その判断基準は体系的に文書化されておらず、フィードバックコメント（feedback.md）と確信度スコア（phase2/）の中に暗黙的に散在している状態であった。本プロジェクトは、この散在する暗黙知を体系的に抽出・文書化し、最終的には AI エージェントとして自律的に批評・判断できるシステムを構築することを目標とする。

従来の運用プロセスとAIの限界

プロジェクト開始以前から、チームでは AI を分析補助に活用していたが、いくつかの本質的な限界が明らかになっていた。AI は入力されたアナリストレポートから「競争優位性らしき部分を抽出して競争優位としている」（FM Y の COST 全体印象コメントより）に留まり、真に構造的な優位性と表面的な強みを区別する能力を欠いていた。さらに、AI の推論が銘柄ごとに不安定であること、推論過程で生じた批判的視点が最終出力で消失すること、そして何より FM Y 自身が最も望む「既存判断への警鐘機能」を提供できていなかったことが、本プロジェクトの出発点となった。

1.2 ファンドマネージャー Y の投資判断プロセス

Y の確信度スコアリング体系

FM Y は、AI が生成した競争優位性の各項目に対して、5段階の確信度スコアを直接付与する体系を持つ。これは Phase 2 スコアリングと呼ばれ、本プロジェクトにおいて「最も信頼度の高いデータ」と位置づけられている。

ランク	確信度	判断基準
かなり納得	90%	構造的優位性が明確 + CAGR への接続メカニズムが具体的 + 定量的裏付けあり
おおむね納得	70%	妥当な仮説として成立 + 一定の裏付けが存在
まあ納得	50%	方向性は認めるが、裏付けが不十分
あまり納得しない	30%	飛躍的解釈が含まれる、または因果関係が逆転している
却下	10%	事実誤認がある、または競争優位性として成立しない

この5段階は単なる評価尺度ではなく、FM Y の投資哲学そのものを反映している。最高ランク（90%）の条件である「構造的優位性 + 明確な CAGR 接続 + 定量的裏付け」という三要素は、Y が投資判断において何を最も重視するかを端的に示している。

Y が批評・フィードバックを行うフロー

FM Y の批評プロセスは二層構造をとる。第一層は定性フィードバック（feedback.md）であり、AI 出力の各項目に対して具体的な修正指示や批判コメントを記述する。第二層は確信度の直接スコアリング（phase2/）であり、各競争優位性項目に 10%～90% のスコアを付与する。この二層構造により、「なぜその評価なのか」（定性）と「どの程度納得しているのか」（定量）の両面から FM Y の判断を捕捉できる設計となっている。

Phase 1 から Phase 2 への流れ

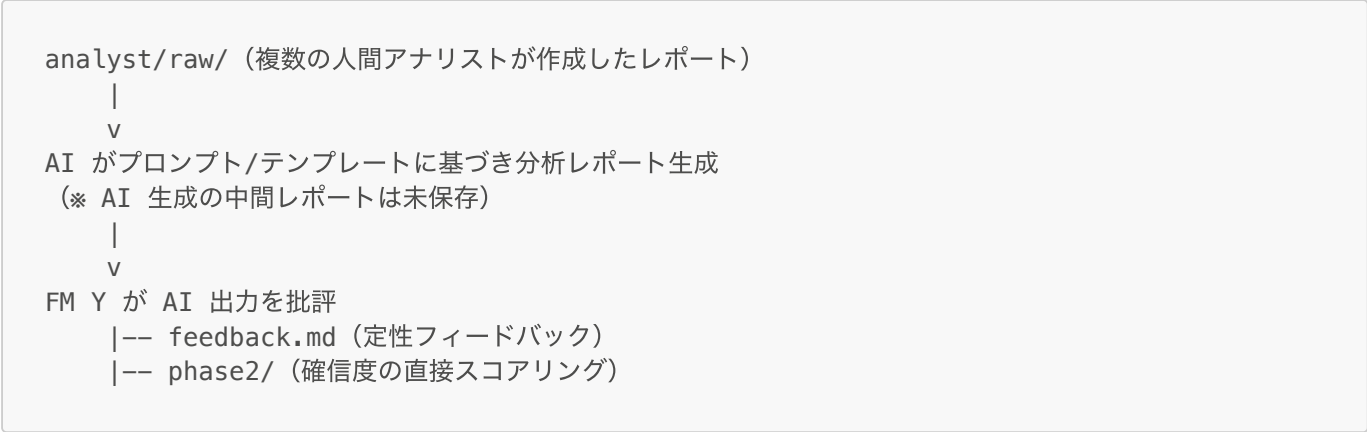
分析プロセスは Phase 1（AI 生成レポート）と Phase 2（確信度スコアリング）の二段階で構成される。Phase 1 では、人間アナリストが作成したレポート（analyst/raw/）を入力として、AI がプロンプトテンプレート（template\_ver2.md）に基づき競争

優位性分析レポートを生成する。Phase 2 では、FM Y がこの AI 生成レポートを読み、各項目に対して確信度スコアと定性コメントを付与する。この Phase 2 の出力こそが、FM Y の暗黙知を最も濃縮して含むデータであり、本プロジェクトの最重要資産である。

### 1.3 現行の AI 分析システム（Phase 1/2）の概要

#### 分析フローの全体像

現行システムの分析フローは以下の通りである。



このフローにおいて注意すべき点は、AI が生成した中間レポートが保存されていないことである。これは、FM Y の批判内容から「何がダメだったか」を逆推定する必要があることを意味し、暗黙知抽出の難度を高めている要因の一つとなっている。

#### Dogma.md と AI レポートの関係

プロジェクト初期の議論において重要な誤認が修正された。**Competitive\_Advantage\_Dogma.md** はアナリスト K の投資判断軸であり、FM Y の投資哲学ではない。さらに、FM Y が批評した AI レポートは、この Dogma.md に基づいて生成されたものではなく、別のプロンプト・テンプレート（template\_ver2.md）に基づくものであった。Dogma.md はサンプルとして後から配置されたものであり、現行の AI 分析フローとは独立した位置づけにある。この誤認の修正は、プロジェクトの方向性を正しく定めるうえで決定的に重要であった。

#### 現在のシステムの限界と課題

現行システムには、AI の推論が銘柄ごとに不安定であること、Few-shot examples（MSFT/TSMC のフォーマット例）の影響で AI がフォーマット模倣に引きずられること、推論段階で生じた批判が最終出力（STEP3）で消失すること、そして FM Y の投資哲学が未文書化であるために AI のプロンプトに組み込めないことなど、構造的な課題が存在する。これらの課題は Session 1 の議論を通じて体系的に特定され、改善計画の基盤となった。

### 1.4 既存資産の棚卸し

本プロジェクトが活用可能な既存資産は以下の通りである。これらは FM Y の暗黙知を抽出するための原材料であり、AI エージェント構築の基盤となる。

#	資産	パス	性質・用途
---	----	----	-------

#	資産	パス	性質・用途
1	アナリスト K の判断軸	<a href="#">analyst/Competitive_Advantage/Competitive_Advantage_Dogma.md</a>	K の投資哲学。サンプル・参考用。 Y 版 Dogma 作成時のフォーマット参考
2	Y の定性フィードバック	<a href="#">analyst/Competitive_Advantage/analyst_YK/feedback.md</a>	AI 出力への修正・批判。Y の暗黙知の宝庫
3	Y の直接スコアリング	<a href="#">analyst/phase2_KY/</a>	5銘柄 (CHD, COST, LLY, MNST, ORLY) の確信度評価。最も信頼度の高いデータ
4	アナリストレポート	<a href="#">analyst/raw/</a>	12銘柄の詳細レポート。AI 分析の入力データ
5	AI 生成用プロンプト	<a href="#">analyst/Competitive_Advantage/Competitive_Advantage_template_ver2.md</a>	AI がレポートを生成する際に使用したテンプレート
6	フォーマット例	feedback.md 内の MSFT/TSMC セクション	AI が参照するフォーマット規範。 Few-shot examples として機能

#	資産	パス	性質・用途
7	Powers テンプレート	analyst/prompt/seven_powers_extract.md	Hamilton Helmer の7 Powers に基づく JSON 抽出 テンプレート
8	Phase 1 サンプル	analyst/phase1/ANET_phase1.md	AI 生成レ ポートの サンプル

これらの資産のうち、項目 2（定性フィードバック）と項目 3（直接スコアリング）が FM Y の暗黙知を抽出するための最重要データである。特に Phase 2 のスコアリングは「FM Y が直接付けたもの」であり、中間に AI や他者の解釈が介在しない、最も純度の高い判断データとして位置づけられる。

## 第2部: FM Y の投資哲学の抽出と体系化

### 2.1 Phase 2 スコアリングパターンの横断分析

#### 5銘柄の全体傾向

FM Y の Phase 2 スコアリングは、5銘柄（CHD: Church & Dwight、COST: Costco Wholesale、LLY: Eli Lilly、MNST: Monster Beverage、ORLY: O'Reilly Automotive）を対象に実施された。これらはいずれも米国上場の優良企業であり、各社が持つとされる競争優位性の各項目に対して、FM Y が 10%～90% の確信度スコアを付与している。

横断分析から浮かび上がる全体傾向として、以下の3点が挙げられる。

第一に、FM Y は「構造的な優位性」と「表面的な強み」を峻別する。市場シェアの高さやブランド認知度といった「結果」を競争優位性として提示する分析には低いスコアを付与し、その結果をもたらした「メカニズム」や「能力」を示す分析には高いスコアを付与する。

第二に、定量的な裏付けの有無がスコアを大きく左右する。同じ方向性の分析であっても、具体的な数値（退職率、会員更新率、CAGR 寄与度等）を伴う分析は高評価となり、定性的な記述に留まる分析は低評価となる。

第三に、FM Y は業界横断的な比較視点を重視する。「メガファーマなら誰でも持っている」「多くの企業が言っていること」といったコメントに見られるように、分析対象企業に固有の優位性であるかどうかを厳格に問う姿勢が一貫している。

#### スコアリング体系の詳細

FM Y のスコアリング体系は、投資判断における「確信度」を5段階で表現するものであるが、各段階の境界は単なる程度の差ではなく、質的に異なる判断基準に基づいている。

**90%（かなり納得）** は、分析が構造的優位性を正確に捉えており、その優位性が企業の CAGR（Compound Annual Growth Rate: 年平均成長率）にどのように接続するかメカニズムが明確で、かつ定量的な裏付けを伴う場合に付与される。例えば、COST の顧客粘着力に関する分析で「上級会員ミックス改善 → ARPU（Average Revenue Per User: 顧客単価）向上 → 既存店売上成長に直接影響」という因果チェーンが示された項目に 90% が付与されている。

**70%（おおむね納得）** は、仮説として妥当であり、一定の裏付けがある場合に付与される。CHD の買収能力に関する分析で「買収基準の明確化 → 人員投入・精査 → 選定能力」という因果チェーンが示された項目がこれに該当する。

**50%（まあ納得）** は、方向性は認めるが裏付けが不十分な場合に付与される。分析の着眼点は評価するが、説得力のある証拠が不足している状態を示す。

**30%（あまり納得しない）** は、飛躍的な解釈や因果関係の逆転が見られる場合に付与される。MNST の「低所得 → コンビニ購買率高い → 粘着力高い」のような論理の飛躍が典型例である。

**10%（却下）** は、事実誤認が含まれるか、そもそも競争優位性として成立しない分析に付与される。MNST の「OPM を GPM と誤認」のような明らかな事実誤りがこれに該当する。

## 2.2 Y が却下する典型パターン（10-30%）7類型

Phase 2 のスコアリングデータとコメントの横断分析から、FM Y が低評価（10-30%）を付与する典型的なパターンが7類型として特定された。これらは「FM Y が競争優位性として認めないもの」のリストであり、AI エージェントが回避すべき分析パターンを定義する。

### 類型1: 結果を原因と取り違い

FM Y が最も頻繁に、かつ最も強い口調で指摘するパターンである。MNST の Phase 2 評価 #1 において、高い市場シェアを競争優位性として提示した分析に対し、FM Y は「高シェアは結果であり背後要因が乏しい」と批判した。同様に、CHD の分析においても「ポートシフトは結果、管理力が優位性」と指摘している。FM Y の哲学において、観察可能な成果（市場シェア、売上成長率等）はあくまで「結果」であり、それをもたらした構造的な「能力」や「メカニズム」こそが真の競争優位性である。

### 類型2: どの企業にも言える一般論

COST の Phase 2 評価 #5 において、分析が提示した優位性に対して FM Y は「多くの企業が言っていること」と一蹴した。競争優位性の本質は差別化にあり、業界の多くの企業が同様に主張できる内容は、たとえ事実であっても優位性としては認められない。FM Y はこの判断基準を feedback.md においても「メガファーマなら誰でも持っている」という表現で繰り返し強調しており、分析対象企業に固有の要素であることを厳格に求めている。

### 類型3: 競合との差別化が不明

COST の Phase 2 評価 #3 において、FM Y は「純粹競合に対して優位な設計かは不明」と指摘した。ある企業の戦略や施策が優れていても、競合企業と比較した際の差別化要因が明示されなければ、競争優位性としては不十分であるという判断基準を反映している。FM Y は「優位性は相対的概念」という原則を一貫して適用している。

### 類型4: 事実誤認

MNST の Phase 2 評価 #6 において、AI が OPM（Operating Profit Margin: 営業利益率）を GPM（Gross Profit Margin: 粗利益率）と誤認した分析に対して、FM Y は即座に却下（10%）を付与した。財務データの正確性は分析の前提条件であり、事実誤認を含む分析は内容の如何にかかわらず信頼に値しないという厳格な基準を示している。

### 類型5: 飛躍的解釈

MNST の Phase 2 評価 #5 において、「低所得層はコンビニでの購買率が高い → Monster の主要販売チャネルはコンビニ → よって顧客粘着力が高い」という論理展開に対し、FM Y は「飛躍」と断じた。各ステップの間に十分な論理的接続がなく、前提から結論への推論に論理的な跳躍がある場合、FM Y はそれを明確に否定する。

### 類型6: 定性的で定量的裏付けなし

COST の Phase 2 評価 #2 において、売り場づくりに関する分析に対して FM Y は「定性的差異がわかりにくく定量評価もしがたい」と指摘した。FM Y の評価体系において、定量的な裏付けは確信度を高める最も重要な要素の一つである。方向性が正しく

ても、数値的な根拠を伴わない定性的記述のみでは、投資判断に必要な確信度に到達しない。

## 類型7: 戦略を優位性と混同

ORLY の Phase 2 評価において、「ドミナント出店戦略」を競争優位性として提示した分析に対し、FM Y は「戦略は優位性ではないのでは」と疑問を呈した。戦略（Strategy）は企業が選択する行動方針であり、競合も同様の戦略を採用し得る。一方、優位性（Competitive Advantage）は戦略の実行を通じて構築される、競合が容易に模倣できない構造的な強みである。この区別は FM Y の判断体系において根本的な原則として機能している。

## 2.3 Y が高く評価する典型パターン（70-90%）6類型

FM Y が高評価（70-90%）を付与する分析パターンは、却下パターンの裏返しではなく、独自の積極的な評価基準を持つ。以下の6類型は「FM Y が競争優位性として認める条件」を定義する。

### 類型1: 定量的裏付け

COST の Phase 2 評価 #1 において、従業員の定着率に関する分析が「退職率 9% vs 業界平均 20%+」という具体的な数値比較を提示したことに対し、FM Y は「意味のある数値裏付けは納得感を高める」と評価した。ここで重要なのは、単に数値が示されているだけでなく、その数値が「意味のある」（=競争優位性の存在を直接的に示唆する）ものであることが求められる点である。形式的な数値の羅列ではなく、分析の主張を実質的に裏付ける数値が高評価の条件となる。

### 類型2: 直接的な CAGR 接続メカニズム

COST の顧客粘着力に関する分析で、FM Y が最高評価（90%）を付与した項目は、「上級会員ミックス改善 → ARPU 向上 → 既存店売上成長に直接影響」という明確な CAGR 接続メカニズムを示していた。FM Y は競争優位性を単独で評価するのではなく、その優位性が企業の持続的成長（CAGR）にどのように寄与するかのメカニズムを重視する。抽象的な「成長に寄与する」という記述ではなく、因果の連鎖を具体的に示すことが求められる。

### 類型3: 能力 > 結果

CHD の Phase 2 評価 #3 において、FM Y は「ポート管理能力が優位性、ポートシフトは結果」と明確に述べた。これは類型 1（却下パターン: 結果を原因と取り違え）の対極に位置する評価基準である。FM Y は、企業が持つ「能力」（Capability）を「結果」（Outcome）よりも上位の概念として評価する。CHD の事例では、ブランドポートフォリオのシフト（結果）ではなく、それを実現するポートフォリオ管理能力（能力）が真の競争優位性であると判断された。

### 類型4: 明確な因果チェーン

CHD の Phase 2 評価 #1 において、買収能力に関する分析が「買収基準の明確化 → 人員投入・精査 → 選定能力」という因果チェーンを示したことに対し、FM Y は 70% の確信度を付与した。各ステップが論理的に接続され、前段が後段の前提条件となる明確な因果連鎖の提示は、FM Y の評価を高める重要な要素である。飛躍的解釈（却下類型5）の対極として位置づけられる。

### 類型5: 他社事例との類推

MNST の Phase 2 評価 #1 の詳細コメントにおいて、FM Y はブランド力の評価に際して、ロレアル、ロリンス、アトラスコブコ、MSFT、AAPL、Disney、エルメス、フェラーリなどの過去の判断事例を明示的に参照した。FM Y は個別銘柄の分析を孤立に行うのではなく、過去に評価した企業群との比較・類推を通じて判断の一貫性を担保している。この類推能力は FM Y の暗黙知の中でも特に体系化が重要な要素である。

### 類型6: 構造的 vs 補完的の区別

COST のガソリン戦略に関する分析において、FM Y は CAGR への寄与を「構造的要素」と「補完的要素」に分類する視点を示した。構造的要素とは企業の持続的成長の根幹をなす優位性であり、補完的要素とはそれを支援・強化する副次的な要因である。



FM Y はこの区別を通じて、投資判断における優先順位付けの基準を明確にしている。全ての優位性が同等ではなく、CAGR への構造的な寄与度に基づいた序列化が求められる。

---

## 2.4 Y のブランド力評価フレームワーク（4類型）

FM Y は MNST の Phase 2 コメントにおいて、ブランド力を評価するための体系的なフレームワークを展開した。注目すべきは、FM Y が過去に分析した複数の企業事例を明示的に引用しながら類型化を行っている点であり、これは長年の投資判断経験に基づく高度に体系化された暗黙知の発露である。

### 類型1: ブランド「ポートフォリオ」運営力

企業が複数のブランドを戦略的に取得・管理する能力に基づく競争優位性である。FM Y はロレアル（化粧品ブランドの買収・ポートフォリオ運営）、ロリンズ（害虫駆除サービスのブランド統合）、アトラスコプロ（地域別にマルチブランド/単独ブランド戦略を使い分け）を具体例として挙げた。この類型の本質は、個々のブランドの知名度ではなく、ブランド群を管理・最適化する「経営能力」にある。

### 類型2: 知名度・ブランド認知

圧倒的な知名度や認知度そのものが参入障壁となるケースである。FM Y は3つのサブカテゴリを区別している。第一に、圧倒的知名度（MSFT、AAPL、AMZN、Disney）。第二に、デファクトスタンダードとしての認知度（MSCI、Moody's、S&P）。第三に、差別化された認知（SBUX: スターバックス）。それぞれブランド認知の「質」が異なり、投資判断への含意も異なる。

### 類型3: 価値の維持・通貨代替機能

製品やブランドが時間経過に伴う価値減耗に抵抗する、あるいは価値保存手段として機能するケースである。FM Y は高残価価値を維持するブランド（AAPL、エルメス、フェラーリ）と、通貨の代替的な価値保存機能を持つブランド（マルボロ、ディアジオ）を区別した。この類型は、ブランドの「耐久性」という次元での評価基準を提供する。

### 類型4: ブランドキーパーソンの経営層取り込み

デザインやブランド管理の責任者を C-suite（最高経営幹部）レベルに配置することで、ブランド価値を経営の中枢に組み込むアプローチである。コーチ/タペストリー、AAPL がその例として挙げられた。

ただし、FM Y は自らのブランド力評価フレームワーク全体に対して、「外見的な評価が多く、背後要因をもとにした評価が少ない印象」という自己批評を加えている。この自己批評は、FM Y が自身の判断基準をも常に批判的に検証する姿勢を示しており、投資哲学の核心である「表面的な結果ではなく構造的な原因を問う」態度が、自己評価の場面にも一貫して適用されていることを示す。

---

## 2.5 feedback.md から抽出した暗黙知（8項目）

FM Y の定性フィードバック（feedback.md）には、Phase 2 のスコアリングでは捉えきれない、より深い投資哲学が言語化されている。以下の8項目は、feedback.md のセッション横断分析から抽出された FM Y の暗黙知である。

### 項目1: 優位性は「名詞」で表現される属性

FM Y は競争優位性をブランド力、スイッチングコスト、エコシステム等の「名詞」で表現される属性として捉える。これは、優位性が一時的な行動や状態ではなく、企業に内在する構造的な「もの」とであるという認識を反映している。動詞（「成長している」「拡大している」）ではなく名詞（「ブランド力」「規模の経済」）で表現できることが、真の競争優位性であるかどうかの試金石となる。

### 項目2: 相対的優位性の要求

FM Y は「メガファーマなら誰でも持っている」という表現で、優位性の相対性を繰り返し強調した。ある能力や資源が優位性であるためには、競合他社と比較して差別化された水準にあることが必要条件であり、業界標準レベルの能力はいかに高度であっても「優位性」とは呼べない。この判断基準は、分析の際に常に比較対象を明示することを求める。

### 項目3: CAGR 接続での確度の分離

FM Y は「優位性と CAGR 下支えの評価は独立して行い、優位性が 100% あるものとして、CAGR 接続への納得度評価とする」という方針を明示した。これは、優位性の存在確認と、その優位性が成長率に寄与するメカニズムの評価を、別個の独立した判断として扱うべきであるという原則である。この分離により、「優位性はあるが CAGR への寄与は限定的」「優位性は疑わしいが CAGR への接続メカニズムは明確」といった細分化された判断が可能になる。

### 項目4: 期初レポート > 四半期レビュー

FM Y は、アナリストが作成する期初レポート（初期の投資仮説）と四半期レビュー（その後の更新情報）の間に明確な優先順位を設けている。期初レポート（1）が主であり、四半期レビュー（2）は従である。AI が四半期レビューの内容を拡大解釈して期初の投資仮説を上書きすることは、本来の分析の主従関係を逆転させる行為として批判される。FM Y は「2の積み重ねから拡大解釈するのではなく、1の妥当性を再検討させるような仕組み考えられないのでしょうか？」と述べ、期初仮説の継続的検証こそが AI の本来的役割であるべきだとした。

### 項目5: 既存判断への警鐘機能

FM Y が AI に最も期待する機能は、ファンドマネージャーやアナリストの既存判断に対する「警鐘」である。「FM や AN が既存判断に縛られて、当初の合理性を欠いていることへの警鐘の機能ほしいです」という FM Y の言葉は、AI の役割を「分析の補助」ではなく「判断の監査」として位置づける意図を明確に示している。人間は確認バイアス（Confirmation Bias）や初期判断への固執（Anchoring Bias）に陥りやすく、AI にはこれらの認知バイアスを検出・指摘する役割が期待される。

### 項目6: 推論の一貫性

FM Y は、銘柄ごとに AI の推論パターンが大きく異なることへの不満を表明した。LLY では「考え直す」という批判的推論が続出する一方、他銘柄では穏当な分析に留まるなど、同一の分析フレームワークを使用しているにもかかわらず推論の質と方向性にばらつきが生じていた。FM Y にとって、分析手法の一貫性は結果の信頼性の前提条件であり、銘柄によって推論の厳しさが変動することは受け入れがたい。

### 項目7: AI の推論段階の批判が最終出力で消える

多段階の推論プロセス（STEP1 → STEP2 → STEP3）を経る中で、途中段階で生じた批判的視点が最終出力に反映されない問題が確認された。具体的には、STEP3 での最終出力において、推論段階の批判が消えて好意的に拡大解釈される現象が発生していた。これは AI の「まとめ」傾向（要約時にポジティブバイアスがかかる）に起因すると考えられ、FM Y が求める批判的分析能力の欠如を如実に示す事例である。

### 項目8: フォーマット例の影響

MSFT と TSMC のフォーマット例を参考として AI に与えた結果、AI が他の銘柄に対しても参考例の分析トーンを模倣しようとし、参照例に含まれる批判的な推論が他銘柄の分析にも波及する現象が観察された。Few-shot examples の選択が AI の出力結果を大きく左右するという発見は、AI エージェント設計において参照例の意図的なコントロールが不可欠であることを示唆している。

## 2.6 現行システムの課題（6つの課題の詳細分析）

Session 1 の議論と Phase 2 データの分析を通じて、現行の AI 分析システムには6つの構造的課題が特定された。これらの課題は相互に関連しており、個別の修正ではなく体系的な改善が必要である。

### 課題1: AI 推論の不安定性

**現象:** 銘柄ごとに AI の推論パターンが大きく異なる。LLY の分析では「考え直す」という批判が続出する一方、他銘柄では比較的穏当な分析に留まる。また、STEP3（最終出力）において、推論途中で生じた批判が消失し、好意的に拡大解釈される。

**根本原因:** Few-shot examples（MSFT/TSMC のフォーマット例）の影響で、AI がフォーマットとトーンを模倣しようとする。参照例に含まれる分析パターンが、分析対象銘柄の特性に関わらず適用されてしまう。

**影響範囲:** 全銘柄の分析品質に影響。FM Y の信頼を損なう最も直接的な要因であり、分析結果の比較可能性を阻害する。

## 課題2: コンテキスト理解の限界

**現象:** AI がアナリストの意図を誤読する。具体例として、CDNS の技術革新に関する分析で、顧客サイドの革新と企業サイドの投資必要性を混同した。また、期初レポート（1: 投資仮説の軸）と四半期レビュー（2: 更新情報）を並列に評価し、2の内容を過度に重視する傾向がある。

**根本原因:** AI は入力テキストの文脈的な重み付けを自律的に行えず、全ての情報を等価に扱おうとする。「1が主、2が従」という構造的な優先順位がプロンプトに十分に組み込まれていない。

**影響範囲:** アナリストレポートの解釈精度全般に影響。特に、投資仮説の時間的変遷を正しく追跡する能力を欠くことで、判断の文脈が失われる。

## 課題3: CAGR パラメータの一貫性

**現象:** LLY では確度と寄与を分離して評価する一方、CHD では確度を寄与に取り込んで評価するなど、銘柄間で CAGR パラメータの処理方法が一貫しない。数値の根拠も不透明であり、「+1-2%」といった寄与度の背景が説明されない。

**根本原因:** CAGR パラメータの処理ルールが明示的に定義されておらず、AI が銘柄ごとにアドホックな処理を行っている。FM Y の方針（「優位性と CAGR 下支えの評価は独立して行う」）がプロンプトに反映されていない。

**影響範囲:** 投資判断の定量的基盤に影響。銘柄間での成長率比較が困難になり、ポートフォリオレベルでの判断に支障をきたす。

## 課題4: 既存判断の無批判的受容

**現象:** AI はアナリストレポートの内容を受容し、さらに拡大解釈する傾向がある。FM Y が最も欲しいのは「既存判断への警鐘」だが、現行の AI はその逆の動作（既存判断の肯定・強化）を行う。

**根本原因:** AI の基本的な動作傾向として、入力テキストの主張を否定するよりも肯定・拡張する方向にバイアスがかかる。批判的分析を行うためには、明示的な指示とフレームワークが必要だが、現行のプロンプトにはその仕組みが組み込まれていない。

**影響範囲:** 本プロジェクトの最も本質的な課題。AI が単なる「分析補助ツール」に留まるか、「判断の監査役」として機能するかの分水嶺であり、FM Y の期待する AI の役割を実現できるか否かを左右する。

## 課題5: AI が結果を原因に取り違える

**現象:** Phase 2 で FM Y が繰り返し指摘した問題。AI は「高い市場シェア」「売上成長」などの観察可能な結果を、そのまま競争優位性として提示する。FM Y のコメント:「シェアは結果であって原因ではない」「ポートシフトは結果、管理能力が優位性」。COST の全体印象として「競争優位性らしき部分を抽出して競争優位としている」。

**根本原因:** AI はテキスト中のポジティブな記述を抽出する能力は高いが、因果関係の方向性（原因 → 結果）を識別する能力が不足している。「結果を生み出すメカニズム」を問う思考は、明示的なプロンプト設計なしには実現しにくい。

**影響範囲:** 競争優位性分析の根幹に影響。「何が優位性か」の定義そのものに関わるため、この課題が解決されなければ分析全体の信頼性が担保されない。

## 課題6: Y の哲学が未文書化

**現象:** アナリスト K の投資判断軸 (Dogma.md) は文書として存在するが、FM Y の判断軸は体系化されていない。FM Y の判断パターンは Phase 2 スコアリングと feedback.md に暗黙的に含まれるが、明示的なルール集として抽出・整理されていない。FM Y と K のフレームワークの差分も不明である。

**根本原因:** FM Y の投資哲学の文書化がプロジェクト開始以前には行われておらず、暗黙知のままであった。Phase 2 のスコアリングデータは存在するが、それを体系的な判断基準として再構成する作業が未実施であった。

**影響範囲:** 全ての改善計画の前提条件。FM Y の哲学が文書化されなければ、Critique Agent のシステムプロンプトに組み込むことができず、Stage 2 以降の全てのステップが実行不可能となる。本課題は Stage 1 として最優先で取り組まれ、判断パターン横断分析レポートと Y 版 Dogma ドラフトの作成により完了した。

## 第3部: 学術研究サーベイ --- LLMマルチエージェント投資システムの最前線

### 3.1 調査対象論文の概要

本サーベイでは、LLM (Large Language Model) を中核としたマルチエージェント・システム (Multi-Agent System; MAS) を金融投資に応用する近年の学術研究9本を体系的に整理する。これらの研究は、エージェントの役割設計、議論プロトコル、バイアス排除、ポートフォリオ最適化といった多岐にわたる課題に取り組んでおり、次世代の資産運用アーキテクチャの輪郭を描き出している。

#	タイトル	著者/機関	主要貢献	キーワード
1	AlphaAgents: Large Language Model based Multi-Agents for Equity Portfolio Constructions	BlackRock	3専門エージェントによるラウンドロビン討論と Financial RAG	エージェント討論, RAG, ポートフォリオ構築
2	MASFIN: A Multi-Agent System for Decomposed Financial Reasoning and Forecasting	Rochester Inst. of Tech	5段階パイプラインと生存者バイアス排除	パイプライン分解, Postmortem, ライブ検証
3	TradingAgents: Multi-Agents LLM Financial Trading Framework	UCLA / MIT	Bull vs Bear の弁証法的討論とリスク管理仲裁	弁証法, リスク管理, ファンドマネージャー
4	HedgeAgents: A Balanced-aware Multi-agent Financial Trading System	South China Univ. of Tech	マルチアセット予算配分会議と最適化	バランス運用, 予算配分, 多資産
5	FinRobot: An Open-Source AI Agent Platform for Financial Applications	AI4Finance Foundation	オープンソース基盤とスマートスケジューラ	オープンソース, FinRL, MDP
6	StockBench: Can LLM Agents Trade Stocks Profitably In Real-world Markets?	Tsinghua Univ.	リアル市場検証とデータ汚染チェック	ベンチマーク, Contamination, 実市場
7	Automate Strategy Finding with LLM in Quant Investment	HKUST	Seed Alpha自動生成と動的戦略最適化	Alpha Factory, Cross-Section, MLP
8	Assessing Look-Ahead Bias in Stock Return Predictions Generated By GPT Sentiment Analysis	Glasserman et al.	GPTセンチメント分析における先読みバイアスの定量評価	先読みバイアス, デバイアス, センチメント

#	タイトル	著者/機関	主要貢献	キーワード
9	Can LLM-based Financial Investing Strategies Outperform the Market in Long Run?	Univ. of Edinburgh	FINSABERフレームワークによる長期パフォーマンス検証	長期バックテスト, レジーム分析, FINSABER

以下では、各研究の設計思想、技術的貢献、および実証結果を詳述する。

### 3.2 AlphaAgents (BlackRock)

BlackRockが提案するAlphaAgentsは、LLMベースのマルチエージェント・システムを用いて株式ポートフォリオを自律的に構築するフレームワークである。本研究の最大の独自性は、伝統的なファンドの投資委員会における議論プロセスを、3つの専門エージェント間の構造化された「討論 (Debate)」として再現した点にある。

#### 3つの専門エージェントの役割定義

AlphaAgentsは、投資判断に必要な分析を3つの独立した視点に分解する。

- Fundamental Agent:** 10-K/10-Q等のSEC開示書類を解析し、企業の「本源的価値 (Intrinsic Value)」を評価する。収益性、キャッシュフロー、事業マージンの推移を精査し、定性的なビジネス評価を行う。
- Sentiment Agent:** Bloomberg等のニュースソース、アナリスト評価、インサイダー取引の動向を分析し、「市場心理 (Market Sentiment)」を定量化する。Reflection-enhanced Summarization (反射型要約) と呼ばれる手法を用い、情報の文脈を批判的に解釈する。
- Valuation Agent:** 歴史的な株価と出来高データから、統計的な「価格の妥当性」を判定するクオンツ分析を担う。

#### エージェント別のデータソースとツール構成

エージェント	データソース	使用ツール	主な分析メトリクス
Fundamental	10-K/10-Q (SEC), yfinance API	Financial RAG Tool	営業利益率, キャッシュフロー, 事業進捗
Sentiment	Bloomberg News, アナリスト評価	Reflection-enhanced Summarization Tool	センチメントスコア, 経営陣の安定性
Valuation	Yahoo Finance (株価・出来高)	Volatility & Return Calculation Tool	年率リターン, 年率ボラティリティ, モメンタム

#### Financial RAGツールの詳細

Fundamental Agentが使用するFinancial RAG (Retrieval-Augmented Generation) ツールは、単純なキーワード検索ではなく、SEC開示書類のセマンティック検索を実現する。Arize Phoenix等の評価フレームワークと連携し、検索結果のFaithfulness (誠実性) とRelevance (関連性) を常時監視することで、エージェントがハルシネーションに基づく分析を行うリスクを抑制している。

#### パフォーマンス指標の数式

Valuation Agentは、以下の数式に基づく計算ツールを実装している。

年率換算リターン (Annualized Return) :

$$R_{\text{annualized}} = \left( (1 + R_{\text{cumulative}})^{\frac{252}{n}} \right) - 1$$

年率換算ボラティリティ (Annualized Volatility) :

$$\sigma_{\text{annualized}} = \sigma_{\text{daily}} \times \sqrt{252}$$

ここでnは取引日数、252は米国市場の年間取引日数である。これらの指標により、エージェントは統計的優位性を定量的に報告する。

ラウンドロビン方式の議論プロトコル

AlphaAgentsの中核的イノベーションは、Microsoft AutoGenのインフラを活用したRound Robin方式の討論プロトコルである。各エージェントは順番に自らの分析結果を提示し、他のエージェント（ピア）の分析をレビューした上で、自身の見解を更新・修正する。

Zscaler（ZS）銘柄のケーススタディがこのプロセスを鮮明に示す。Valuation Agentは直近3ヶ月の約50%の株価上昇に基づき「BUY」を推奨した。しかしFundamental Agentが営業利益率-14.5%という赤字体質と、経営幹部（Amit Sinha氏ら）によるインサイダー売却を指摘。Sentiment Agentもニュースの論調がネガティブであることを裏付けた。討論を経て、当初「BUY」を主張したValuation Agentが見解を修正し、最終的に全エージェントが「SELL」でコンセンサスを形成した。この事例は、単一指標への依存がもたらす過剰自信バイアスを、構造化された討論が排除できることを実証している。

3.3 MASFIN（Rochester Inst. of Tech）

MASFINは、投資判断を5つの独立した段階（Crew）に分解し、各段階を専門エージェントチームが担当するシーケンシャル・パイプライン型のアーキテクチャである。この設計の独自性は、投資プロセスの「失敗からの学習」を最初のステージとして明示的に組み込んだ点にある。

5段階パイプラインの詳細

段階	名称	プロセス	成果物
1	Postmortem	上場廃止・倒産企業（NKLA, BBBYQ, RIDE, BLUE等18社）の失敗パターンを分析し、財務的不備・法令違反・センチメント崩壊の3類型に整理	リスク回避フラグ、失敗シグナル
2	Screening	Alpha Factory（9カテゴリのアルファ因子）と失敗パターンフィルタを組み合わせ、数千銘柄から50-100銘柄を選別	投資ユニバース候補リスト
3	Analysis	Fundamental/Sentiment/Valuation の3エージェントが各銘柄を深掘り分析。OHLCV、財務比率、センチメントを統合評価	銘柄別スコア、深層レポート
4	Timing	RSI-14、リターンZスコア、5日移動平均線との乖離、出来高トレンドに基づき最適エントリー時期を判定	BUY/SELL/HOLDアクション
5	Portfolio	15-30銘柄に絞り込み、リスクプロファイル（Risk-Averse / Risk-Neutral）に応じたウェイト最適化を実行	最終ポートフォリオ構成案

各段階の詳細プロセス

**Postmortem Crew**は、2020年から2025年の期間に倒産または上場廃止となった18社のデータを分析する。EVセクターのNikola（NKLA）、小売のBed Bath & Beyond（BBBYQ）、バイオテクのBluebird Bio（BLUE）等が対象であり、これにより生存者バイアス（Survivorship Bias）を構造的に排除する。

**Screening Crew**では、Momentum、Mean Reversion、Volatility、Fundamental、Growth、Liquidity、Quality、Sentiment、Technicalの9カテゴリからアルファ因子を抽出し、カテゴリーバランスを維持しながら候補を選定する。

**Analysis Crew**では、時間的整合性（Temporal Alignment）の維持が厳格に要求される。各エージェントは分析対象日以前に公開されたデータのみアクセスが制限される「Contemporaneous Data Snapshots（同時性のスナップショット）」方式を

採用し、先読みバイアスを物理的に遮断する。

**Timing Crew**では、Sortino Ratio、Zスコア、回帰スロープ等の先行指標を組み合わせ、エントリーの妥当性を多面的に検証する。

**Portfolio Crew**では、リスクプロファイルが投資判断に直接影響する。Risk-Averse設定では高ボラティリティ銘柄が自動的に除外され、Risk-Neutral設定ではモメンタムを重視したポートフォリオが構築される。

### 8週間ライブ検証の実績

MASFINは8週間のライブ運用検証において**7.33%の累積リターン**を達成した。同期間のS&P 500（4.92%）を大幅に上回り、マルチエージェント・パイプラインの実効性を実証した。各Crew間に配置されたSummary Agentが分析報告書を要約し、Human-in-the-Loop（HITL）による検証ポイントとして機能することで、ハルシネーションの蓄積を遮断している点も注目に値する。

## 3.4 TradingAgents（UCLA / MIT）

TradingAgentsは、金融市場の意思決定を「弁証法（Dialectical Method）」の構造で再現するフレームワークである。AlphaAgentsが同一視点の3エージェント間の合意形成を重視するのに対し、TradingAgentsはあえて対立する立場を制度的に設計する点が独自である。

### 強気派（Bull）vs 弱気派（Bear）の弁証法的構造

システムには、意図的に対立する役割が割り当てられる。

- **Bullish Researcher（強気派）**：企業の成長ポテンシャル、エコシステムの競争優位性、財務の健全性を強調し、買いを推奨する立場をとる。
- **Bearish Researcher（弱気派）**：競合に対する劣位、地政学的リスク（サプライチェーンの中国依存等）、バリュエーションの割高感を指摘し、リスクを警告する立場をとる。

Apple（AAPL）株のケーススタディでは、Bullish Researcherがスマートホーム・AI分野への進出と巨大エコシステムの競争力を主張した一方、Bearish ResearcherはAI参入の遅れ、地政学的緊張、PERの割高感を反論として提示した。

### リスク管理チームの仲裁機能

両者の主張を受けたRisk Manager（リスク管理担当）が仲裁者として機能する。Risk Managerは「強気派の言う成長性は認めるが、弱気派の言うボラティリティリスクも無視できない」と判断し、全力買いではなくドル・コスト平均法による段階的購入とストップロス設定という折衷案を提示した。

### ファンドマネージャーによる最終判断

TradingAgentsの最上位にはLead Trader / Fund Managerが位置し、各アナリストの提案を統合して最終的なBuy/Sell/Holdと取引量を決定する。この階層構造により、弁証法的討論の成果が実行可能なアクションに変換される。

## 3.5 HedgeAgents（South China Univ. of Tech）

HedgeAgentsは、株式のみならず複数の資産クラスにまたがるバランス運用を志向するマルチエージェント・システムである。単一資産への集中リスクを回避し、ポートフォリオ全体のエクスポージャーを管理する点に独自性がある。

### バランス重視のマルチアセット運用

HedgeAgentsは、株式、債券、コモディティ等の複数資産クラスを対象とし、セクター間の分散を制度的に強制する。リスクマネージャー・エージェントがポートフォリオ全体の露出を監視し、最大ドローダウンの抑制を最優先の制約条件として設定する。

## 予算配分会議（Budget Allocation Conference）

HedgeAgentsの独自メカニズムは、エージェント間で開催される「予算配分会議（Budget Allocation Conference）」である。各資産クラスを担当するエージェントが、自身の担当領域への配分増加を主張し、リスクマネージャーがポートフォリオ全体の制約条件の下で調停を行う。このプロセスにより、市場環境の変化に応じた動的な資産配分が実現される。

## 最適化目的関数

HedgeAgentsのポートフォリオ最適化は、以下の目的関数に基づく：

$$\begin{aligned} \max_w \quad & \mu^T w - \frac{\lambda}{2} w^T \Sigma w \\ \text{s.t.} \quad & \sum_i w_i = 1, \quad w_i \geq 0 \end{aligned}$$

ここで $w$ はポートフォリオのウェイトベクトル、 $\mu$ は期待リターンベクトル、 $\Sigma$ は共分散行列、 $\lambda$ はリスク回避度パラメータである。Financial Chain-of-Thought（CoT）と階層的メモリ検索（Hierarchical Memory Retrieval）を組み合わせ、過去の投資判断とその結果を教訓として現在の意思決定に活用する。

## 3.6 FinRobot（AI4Finance Foundation）

FinRobotは、金融AI研究のためのオープンソースプラットフォームとして設計された点で、他の研究とは根本的に異なるポジションを占める。個別の投資戦略の提案ではなく、研究者やプラクティショナーが自在にエージェントを構築・評価できる基盤の提供を目的としている。

### オープンソースプラットフォームとしての設計

FinRobotは、エージェントの定義、ツールの統合、バックテスト環境の構築をモジュール化し、研究の再現性を担保する。特に、yfinance、Finnhub等の市場データAPI、RAGツール、テクニカル分析ライブラリとの統合が標準化されている。

### スマートスケジューラとパフォーマンスベースの選抜

FinRobotの技術的特徴は「スマートスケジューラ」にある。複数のエージェントの過去のパフォーマンスを追跡し、市場環境に応じて最も適切なエージェントを動的に選抜・配置する。これにより、強気相場ではモメンタム系エージェント、弱気相場ではディフェンシブ系エージェントが優先的に稼働する適応的運用が実現される。

### 強化学習（FinRL）アプローチ

FinRobotは、強化学習ライブラリFinRLとの統合を提供する。投資行動をマルコフ決定過程（Markov Decision Process; MDP）として定式化し、エージェントが試行錯誤を通じて最適な行動方を学習する枠組みを構築している。

MDPの定式化は以下のとおりである：

$$\langle S, A, P, R, \gamma \rangle$$

- $S$ : 状態空間（ポートフォリオの保有状況、市場指標、テクニカル指標）
- $A$ : 行動空間（各銘柄のBuy/Sell/Hold、取引量）
- $P$ : 状態遷移確率（市場のダイナミクス）
- $R$ : 報酬関数（リスク調整後リターン）
- $\gamma$ : 割引率（将来の報酬の現在価値への変換係数）

エージェントは方策 $\pi(a|s)$ を学習し、期待累積報酬 $\mathbb{E} \left[ \sum_{t=0}^T \gamma^t R(s_t, a_t) \right]$ の最大化を目指す。



### 3.7 StockBench (Tsinghua Univ.)

清華大学のStockBenchは、LLMエージェントが「実際のリアル市場で利益を上げられるのか」という根本的な問いに対し、厳格なベンチマーク環境を提供する研究である。

#### リアル市場でのLLMエージェント検証

StockBenchは、シミュレーション環境ではなく、実際の市場データと執行条件（スリッページ、取引手数料等）を反映した環境でエージェントのパフォーマンスを測定する。これにより、学術研究で報告されるバックテスト上の好成績が、実運用においても再現可能かを検証する。

#### データ汚染 (Contamination) チェック手法

StockBenchの最も重要な技術的貢献は、LLMのデータ汚染チェック手法の提案である。GPT-4o等のLLMは膨大なテキストコーパスで事前学習されており、特定の期間の株価推移や企業イベントを「記憶」している可能性がある。この記憶がバックテスト結果を汚染するリスクに対し、StockBenchは以下のプロトコルを導入した。

- 記憶テスト:** 特定期間の株価推移をLLMに質問し、正確に言い当てられるかを検証する。
- 汚染期間の除外:** 正答率が統計的に有意に高い期間は、バックテスト評価から除外する。
- カットオフ以降のデータ活用:** LLMのトレーニングカットオフ日以降のデータを優先的に検証対象とする。

この手法により、LLMの「未来知識」に起因する見かけ上の好パフォーマンスを排除し、エージェントの真の推論能力を評価することが可能になる。

### 3.8 Automate Strategy Finding (HKUST)

香港科技大学のAutomate Strategy Finding (ASF) は、LLMを用いてクオンツ投資のアルファ・ファクター (Seed Alpha) を自動生成する手法を提案する。人間のクオンツ・リサーチャーが数週間かけて行うファクター探索を、LLMにより大幅に効率化・拡張する点が独自の貢献である。

#### LLMによるSeed Alpha自動生成

ASFでは、LLMに対して11本の主要な金融文献を入力し、合計100個のSeed Alpha (アルファの種) を自動生成させる。生成されたアルファ候補は、Predictive Quality (予測品質) とCategory Balance (カテゴリーバランス) の2軸で評価・フィルタリングされる。

#### Cross-Section演算子とTime-Series演算子

生成されたアルファは、2種類の演算子によって構造化される。

- Cross-Section Operators ( $f_{cs}$ ):** 同一時点における銘柄間の相対的関係を捉える。セクター内の割安度、同業他社比較でのモメンタムの強さ等を定量化する。
- Time-Series Operators ( $f_{ts}$ ):** 個別銘柄の時間的パターンを解析する。過去のリターン推移、ボラティリティの構造変化、移動平均からの乖離等を捕捉する。

#### アルファ統合数式

これら2種類の演算子を統合したシグナルは以下の数式で表現される：

$$\alpha_{ij}^{(t)} = w_{cs} \cdot f_{cs}(X_i^{(t)}) + w_{ts} \cdot f_{ts}(X_i^{(t)}, \dots, X_i^{(t-n)})$$

ここで $\alpha_{ij}^{(t)}$ は時点 $t$ における銘柄 $i$ のアルファ因子 $j$ のシグナル値、 $w_{cs}$ と $w_{ts}$ はそれぞれの演算子の重みである。

#### SSE50における実績

ASFのバックテストはSSE 50（上海証券取引所50指数）を対象とし、2023年1月から2024年1月の期間において**53.17%の累積リターン**を達成した。同期間のベンチマーク（SSE 50指数）が-11.73%と下落する中でのこの成績は、LLMによるアルファ生成の有効性を強く示唆している。

さらに、LLMによる選別後のInformation Coefficient（IC）は顕著な改善を示した。例えば、Momentum因子の平均ICは選別前の0.0092から選別後の0.0208へと倍増し、Volatility、Fundamental等のカテゴリーにおいても同様の改善が確認されている。

### 3層MLPによるDynamic Strategy Optimization

選定されたアルファ群のウェイト最適化には、3層のMLP（多層パーセプトロン）が採用される。

- **入力層:** デイリーのアルファ値
- **中間層:** 10個のReLU活性化ノード
- **出力層:** 予測利回り

この構造により、市場環境（Market Regime）の変化に応じてアルファ因子の重みが動的に調整される。強気相場ではモメンタム系指標（RSI, MACD）の重みが増大し、不確実性の高い相場ではボラティリティ系指標（ATR, Bollinger Bands）やファンダメンタル指標へとウェイトがシフトする。訓練・検証・テストの各フェーズは時間軸で完全に分離され、モデルが将来のボラティリティ構造を学習することを防止している。

---

## 3.9 先読みバイアス研究（Glasserman et al.）

Glassermanらの研究は、GPTを用いたセンチメント分析が内包する「先読みバイアス（Look-Ahead Bias）」の問題を正面から取り上げた重要な研究である。

### GPTセンチメント分析における先読みバイアスの評価

LLMは膨大なテキストデータで事前学習されており、特定の企業イベント（決算発表、倒産、M&A等）の結果を事前に「知っている」可能性がある。この知識がセンチメント分析に混入すると、バックテスト上では優れた予測精度を示すが、実運用ではアルファが消失するという深刻な問題が生じる。

Glassermanらは、この問題を定量的に評価するための方法論を確立した。具体的には、ニュースの見出しや本文に含まれる「結果を示唆する表現」（例：「株価急騰後の...」「決算を受けて下落した...」）が、センチメントスコアに与える影響を統計的に測定した。

### デバイアス処理手法

先読みバイアスを排除するためのデバイアス処理として、以下の手法が提案されている。

1. **時間的隔離（Temporal Isolation）**：バックテストの各時点において、その日以前に発行されたニュースやデータのみをプロンプトに含める。
2. **表現の中立化**：ニュースの見出しから結果を示唆する単語を、LLMを用いて中立的な表現に書き換えてから分析に供する。
3. **メタインストラクション**：エージェントのシステムプロンプトに「あなたは〇〇年〇月〇日にいる。将来の知識を推論に使用してはならない」という時間軸制約を明示的に埋め込む。

これらの手法は、AlphaAgentsやMASFINの議論プロトコルにおいても適用可能であり、MAS全体のバイアス耐性を高める基盤技術として位置づけられる。

---

## 3.10 長期パフォーマンス研究（Edinburgh）

エディンバラ大学の研究は、「LLMベースの投資戦略は長期的に市場を上回れるのか」という実務上最も重要な問いに挑む。多くの先行研究が短期間（数週間～数ヶ月）のバックテストに依拠しているのに対し、本研究は5～10年スパンの長期検証の必要性を強く主張する。

FINSABERフレームワーク

本研究が提案するFINSABER（Financial Strategy Assessment and Backtesting for Evaluating Robustness）フレームワークは、以下の3つの評価軸を統合した長期バックテスト環境である。

- 1. **時間的頑健性（Temporal Robustness）**：強気相場、弱気相場、レンジ相場等、異なる市場レジームにわたる戦略の安定性を検証する。Rolling Sharpe Ratioを4ヶ月～1年のウィンドウで測定し、特定期間の「ラッキーパンチ」を排除する。
- 2. **銘柄ユニバースの広さ**：100銘柄以上を対象とし、セクター偏りのない評価を行う。
- 3. **市場レジーム分析**：強気相場で過度に保守的でないか、弱気相場で過度に攻撃的でないかを、パッシブなインデックス（S&P 500等）との比較で評価する。

長期バックテストの重要性

FINSABERの検証が明らかにした重要な知見は、短期で好成績を示す戦略の多くが、長期的には市場平均に回帰するか、特定のレジームでのみ有効であるという事実である。AlphaAgentsのH1 2024実験では、Risk-Averse（リスク回避型）エージェントが上昇相場においてボラティリティの高いテクノロジー銘柄を除外した結果、ベンチマークを下回った。FINSABERフレームワークは、この結果を「失敗」ではなく「意図したリスク特性の反映」として解釈できる透明性を提供する。

長期バックテストにおける評価指標として、Sharpe Ratio（リスク単位あたりの超過収益）、Sortino Ratio（下方リスクのみを考慮したリスク調整後リターン）、Calmar Ratio（年率リターン / 最大ドロウダウン）、Maximum Drawdown（下落局面での耐性）が体系的に活用される。これらの指標を市場レジーム別に分解することで、戦略の真の頑健性が評価される。

以上の9本の研究は、LLMマルチエージェント投資システムの異なる側面---エージェント設計、討論プロトコル、パイプライン構造、アルファ生成、バイアス排除、長期評価---を相補的にカバーしている。AlphaAgentsとTradingAgentsが示す討論メカニズム、MASFINが体系化したパイプライン・ガバナンス、ASFが実証したアルファ自動生成、そしてGlassermanらとFINSABERが提起するバイアス・長期検証の課題は、いずれも次世代の自律型資産運用システムの設計において不可欠な構成要素である。これらの知見を統合することで、透明性・再現性・頑健性を兼ね備えた投資意思決定フレームワークの構築が可能になる。

第4部: マルチエージェント・システムアーキテクチャの設計

投資判断の質を構造的に転換するためには、個々のAIモデルの性能向上だけでは不十分である。異なる専門性を持つ複数のエージェントが、協調と対立を通じて意思決定を洗練するアーキテクチャ設計が不可欠となる。本章では、AlphaAgents、TradingAgents、HedgeAgents、FinRobotといった先行研究の知見を統合し、マルチエージェント・システム（MAS）のチーム構成、議論メカニズム、合意形成プロトコルを体系的に設計する。

4.1 エージェントチーム構成と役割定義

MASの中核は、役割定義（Role Prompting）された専門エージェント群である。各エージェントは独自のデータソース、分析ツール、評価メトリクスを有し、単一視点への依存を構造的に排除する。以下に、6つの主要エージェントの役割を定義する。

4.1.1 ファンダメンタル・エージェント（Fundamental Agent）

ファンダメンタル・エージェントは、企業の「本源的価値（Intrinsic Value）」を財務データから抽出する役割を担う。AlphaAgentsのシステムプロンプトでは「あなたはバリュエーション・アナリストです。最新の10-K/10-QをRAGツールで解析し、キャッシュフロー、営業利益率、将来目標への進捗を評価せよ。分析はRAGから得られた証拠のみに基づくこと」と定義されている。

項目	内容
役割	財務諸表（損益計算書、貸借対照表、キャッシュフロー計算書）の解析。企業の収益性・健全性・成長性の評価

項目	内容
データソース	10-K/10-Q（SEC EDGAR）、yfinance API
使用ツール	Financial RAG Tool（検索拡張生成による財務報告書の精密検索）
主な分析メトリクス	営業利益率、フリーキャッシュフロー、事業マージン推移、経営陣の定性的目標への進捗

Financial RAGツールの採用は、単なるテキスト要約ではなく、10-Kの特定セクション（リスク要因、MD&A等）から証拠に基づく分析を行う点に特徴がある。これにより、ハルシネーション（もっともらしい虚偽情報の生成）を抑制し、財務分析の信頼性を担保する。加えて、Arize Phoenixなどの評価ツールによるRAG精度の常時監視が組み込まれ、検索品質の劣化を早期に検出する。

4.1.2 センチメント・エージェント（Sentiment Agent）

センチメント・エージェントは、市場参加者の心理的側面を定量化する。AlphaAgentsでは「Bloomberg等のニュース記事を要約し、インサイダー取引の動向やアナリストの格付け変更が市場心理に与える影響を評価し、投資推奨を行え」と定義される。

項目	内容
役割	金融ニュース、SNS、プレスリリースの解析。市場心理と短期的センチメントスコアの算出
データソース	Bloomberg News、アナリスト評価、企業プレスリリース
使用ツール	Reflection-enhanced Summarization Tool（反射型要約ツール）
主な分析メトリクス	センチメントスコア、経営陣の安定性、インサイダー取引動向

反射型要約ツール（Reflection-enhanced Summarization）は、通常的要約に加え、自らの出力を批判的に再評価するプロセスを内蔵している。これにより、ニュースの表面的な論調に過度に引きずられることなく、経営陣の交代やインサイダー取引といった深層のシグナルの文脈を批判的に解析できる。

4.1.3 バリュエーション・エージェント（Valuation Agent）

バリュエーション・エージェントは、「価格の妥当性」を統計的に判定するクオンツ・アナリストの役割を果たす。AlphaAgentsでは「株価と出来高からボラティリティとリターンを数学的に算出せよ。提供された計算ツールを必ず使用し、統計的優位性を報告せよ」と指示される。

中核となる計算式は以下の通りである。

年率換算リターン（Annualized Return）：

$$R_{\text{annualized}} = \left( (1 + R_{\text{cumulative}})^{\frac{252}{n}} \right) - 1$$

年率換算ボラティリティ（Annualized Volatility）：

$$\sigma_{\text{annualized}} = \sigma_{\text{daily}} \times \sqrt{252}$$

ここでnは取引日数、R<sub>cumulative</sub>は累積リターン、σ<sub>daily</sub>は日次リターンの標準偏差である。252は1年間の標準的な取引日数に基づく。

バリュエーション・エージェントは、これらの定量指標に基づきモメンタム（上昇トレンドの強度）やボラティリティ水準を報告し、エントリーポイントの妥当性を統計的に検証する。

4.1.4 マクロ・レジーム・エージェント（Bull/Bear Researchers）

TradingAgentsの設計に基づき、マクロ・レジーム・エージェントは市場全体の環境認識を担当する。特徴的なのは、このエージェントが意図的に「強気派（Bullish Researcher）」と「弱気派（Bearish Researcher）」の2つのサブエージェントに分割される点である。

サブエージェント	役割
強気派（Bull）	金利低下、財政拡大、技術革新といった市場を上昇させる要因を積極的に探索
弱気派（Bear）	地政学リスク、金融引き締め、サプライチェーン途絶といったダウンサイドリスクを探索

この構造的な対立により、マクロ環境の評価が一方向に偏ることを防止する。強気派と弱気派の議論を経て「現在の市場レジーム（Bull Market / Bear Market / Transition）」が判定され、後続のポートフォリオ構築における資産配分の基本方針が決定される。

4.1.5 リスクマネージャー（Risk Manager）

HedgeAgentsの設計に基づくリスクマネージャーは、個別銘柄ではなくポートフォリオ全体のリスク管理を統括する。

項目	内容
主要責務	ポートフォリオ全体のエクスポージャー管理、最大ドローダウンの抑制、セクター間分散の強制
介入トリガー	ボラティリティ急騰、特定セクターへの過度な集中、ドローダウン閾値の超過
出力	リスク調整後の制約条件（最大ウェイト制限、セクター配分上限、損切りライン等）

リスクマネージャーは、アナリストエージェント群の「買い推奨」に対して構造的に異議を申し立てる権限を持つ。Apple（AAPL）のケーススタディ（後述）では、強気派の成長性評価と弱気派のボラティリティ懸念を仲裁し、「段階的購入（ドル・コスト平均法）」と「ストップロス設定」という条件付き投資プランを提示した。

4.1.6 ファンドマネージャー（リードトレーダー）

ファンドマネージャーは、各専門エージェントの分析を統合し、最終的な売買判断（Buy / Sell / Hold）と取引量を決定する最上位エージェントである。

TradingAgentsとHedgeAgentsでは、このエージェントに最終決定権が付与される。ファンドマネージャーは、各アナリストの提案を単純に集約するのではなく、議論の文脈（どのような論拠が対立し、どう解消されたか）を踏まえて統合的判断を下す。HedgeAgentsでは特に、「最適化目的関数（Optimization Objective Function）」を用いて各エージェントの提案を数理的に統合し、リスクとリターンバランスが最適となるポートフォリオ・ウェイトを算出する。

エージェント構成の全体像

エージェント	データソース	使用ツール	主な分析メトリクス
ファンダメンタル	10-K/10-Q (SEC)、yfinance API	Financial RAG Tool	営業利益率、キャッシュフロー、事業進捗
センチメント	Bloomberg News、アナリスト評価	Reflection-enhanced Summarization	センチメントスコア、経営陣安定性
バリュエーション	Yahoo Finance（株価・出来高）	Volatility & Return Calculation Tool	年率リターン、年率ボラティリティ
マクロ・レジーム	FRED、FX、地政学ニュース	Market Regime Classifier	市場レジーム判定（Bull/Bear/Transition）

エージェント	データソース	使用ツール	主な分析メトリクス
リスクマネージャー	ポートフォリオ全体のメトリクス	Risk Constraint Engine	ドローダウン、セクター集中度、VaR
ファンドマネージャー	全エージェントの分析レポート	Optimization Objective Function	統合スコア、最終売買シグナル

## 4.2 エージェント間の議論（Debate）メカニズム

マルチエージェント・システムの本質的価値は、エージェント間の「議論（Debate）」にある。単一モデルの判断では見落とされるリスクや矛盾を、構造化された対話を通じて浮かび上がらせる。先行研究は主に2つの議論メカニズムを提示している。

### 4.2.1 ラウンドロビン方式（AlphaAgents）

AlphaAgentsが採用するラウンドロビン方式は、Microsoft AutoGenのインフラを活用した持ち回り型の議論プロトコルである。

プロセスの詳細：

- 初期分析フェーズ:** 各専門エージェント（ファンダメンタル、センチメント、バリュエーション）が独立に分析を実施し、初期推奨（BUY / SELL / HOLD）を提示
- 相互レビューフェーズ:** 各エージェントがピア（他のエージェント）の分析結果を順次レビューし、自らの分析に対する反論・補強を検討
- 意見更新フェーズ:** 他者の指摘を踏まえて自らの推奨を修正。この更新は合意に至るまで繰り返される

この議論プロセスを支える2つの要素技術がある。

**Financial Chain-of-Thought (Financial CoT)：**金融特化型の思考連鎖推論であり、複雑な投資判断を「データ確認 → 指標算出 → リスク評価 → 推奨決定」の論理的ステップに分解する。これにより、各エージェントの推論過程が透明化され、議論における論拠の追跡が可能となる。

**階層的メモリ検索 (Hierarchical Memory Retrieval)：**過去の投資判断とその結果をメモリとして保持し、類似の市場状況下での教訓を現在の議論に活用する。FinRobotやHedgeAgentsでも同様のメモリ機構が採用されており、エージェントの学習能力と判断の一貫性を担保する。

### 4.2.2 弁証法的アプローチ（TradingAgents）

TradingAgentsは、ラウンドロビンとは異なり、意図的に対立構造を設計する弁証法的アプローチを採用する。

**強気派 vs 弱気派の構造:** マクロ・レジーム・エージェント内に強気派（Bullish Researcher）と弱気派（Bearish Researcher）を配置し、同一銘柄・同一市場環境に対して正反対の論拠を構築させる。これは古典的な弁証法（テーゼ → アンチテーゼ → ジンテーゼ）の構造をエージェントシステムに実装したものであり、確認バイアス（自説に都合の良い情報のみを収集する傾向）を構造的に排除する。

**リスク管理チームの仲裁:** 強気派と弱気派の議論は、リスクマネージャーによって仲裁される。リスクマネージャーは双方の論拠を評価し、ポートフォリオ全体の制約条件と照合した上で、条件付きの統合提案を策定する。

本レポートが目指すシステム（plan.mdに記載のStage 3: マルチエージェント化）では、この弁証法的構造をさらに拡張し、以下の7エージェント構成を採用する：

FM Agent (Y's Philosophy)

├ Bull Agent (買いの根拠を徹底的に探索)

└ Bear Agent (売りの根拠を徹底的に探索)

- Competitive Advantage Critic (判断軸に基づく批判)
- CAGR Validator (数値の根拠と一貫性を検証)
- Causality Checker (結果 vs 原因の区別を強制)
- Historical Analogy Agent (過去判断事例との整合性チェック)
- Bias Detector (確認バイアス・初期判断への固執を検出)

#### 4.2.3 ケーススタディ

##### Zscaler (ZS) : ラウンドロビン議論による合意形成

AlphaAgentsフレームワークにおけるZscaler (ZS) 分析は、議論メカニズムの有効性を実証する典型例である。

##### Step 1 -- 初期分析 (意見の不一致) :

- **Valuation Agent:** ZS株の直近3ヶ月で約50%の上昇トレンドを確認。年率換算リターンは約13.56%でS&P 500を3.85%上回るアウトパフォームを記録。これを根拠に「**BUY**」を推奨
- **Fundamental Agent:** SaaS市場でのリーダーシップを認めつつも、**営業利益率-14.5%** (赤字) とAmit Sinha氏らによるインサイダー株式売却を特定。「**SELL**」を推奨
- **Sentiment Agent:** ニュースの論調がネガティブ化しており、インサイダーの売却が投資家心理に与える負の影響を重視。「**SELL**」を推奨

##### Step 2 -- 議論 (ラウンドロビン) :

各エージェントが他者の分析結果を相互にレビュー。Valuation Agentは、Fundamental Agentが提示した「赤字構造」と「インサイダーの売り」というネガティブな事実を精査し、株価の上昇トレンドがこれらの構造的リスクを正当化するには不十分であると認識を更新した。

##### Step 3 -- 合意 (Consensus) :

「強気トレンドは財務上の懸念とインサイダーの自信欠如を正当化できない」との結論に全員が合意し、\*\*満場一致で「**SELL**」\*\*に決定。当初「**BUY**」を主張していたエージェントが論拠に基づいて意見を変更するこのプロセスこそが、過剰自信バイアスの排除とハルシネーション抑制の核心である。

##### Apple (AAPL) : 弁証法的討論とリスク管理の仲裁

TradingAgentsフレームワークにおけるApple (AAPL) 分析は、弁証法的アプローチの実践例である。

- **Bullish Researcher (強気派) :** 「AppleはスマートホームやAI分野への進出で新たな成長が見込める。独自の巨大なエコシステムがあり、競争力は盤石だ。財務も健全である」と主張し、買いを推奨
- **Bearish Researcher (弱気派) :** 「AI分野への参入はGoogleやAmazonに比べて遅れており、苦戦するだろう。地政学的な緊張 (中国リスク) がサプライチェーンに打撃を与える可能性がある。現在の株価収益率 (PER) は割高だ」と反論
- **Risk Managerの仲裁:** 「強気派の言う成長性は認めるが、弱気派の言うボラティリティリスクも無視できない」と判断。最終的に「段階的購入 (ドル・コスト平均法)」「ストップロス (損切りライン) の設定」という**条件付きの慎重な投資プラン**に修正

このケースは、「買いか売りか」の二択ではなく、条件付きの中間解を導出できる弁証法的アプローチの柔軟性を示している。

## 4.3 合意形成・意思決定メカニズム

エージェント間の議論が収束しない場合、あるいは対立が解消しない場合に、最終判断をどのように下すかは、システムの信頼性に直結する問題である。先行研究は5つの異なるメカニズムを提示している。

### 4.3.1 全会一致型 (AlphaAgents)

ラウンドロビン方式による全会一致型では、全エージェントが合意するまで議論を繰り返す。前述のZscaler事例がこれに該当する。

- **長所:** 判断の堅牢性が最も高い。全員が納得した結論は、特定のエージェントの偏りに影響されにくい
- **短所:** 合意に至るまでの計算コスト（APIコール数）が大きい。市場の急変時に迅速な判断が困難になる可能性がある

4.3.2 マネージャー決定型（TradingAgents / HedgeAgents）

議論の内容を踏まえた上で、最終決定権を持つファンドマネージャー・エージェントが判断を下す。Apple事例のリスクマネージャーによる仲裁もこの変種である。

- **長所:** 迅速な意思決定が可能。責任の所在が明確
- **短所:** マネージャーエージェント自身のバイアスが最終判断に影響するリスク

4.3.3 数理モデルによる最適化と重み付け（HedgeAgents）

エージェントの意見を「A案かB案か」の二択で解決するのではなく、「誰の意見をどの程度採用するか」を数式で決定する。HedgeAgentsの予算配分会議では、マネージャーAIが各エージェントからの期待リターンを評価し、\*\*最適化目的関数（Optimization Objective Function）\*\*を用いてリスクとリターンのバランスが最良となる投資比率を自動算出する。MASFINも同様に、リスク調整後の制約条件に基づく配分比率（Allocation Weights）の最適化を実行する。

- **長所:** 意見の対立を連続値（比率）で解消できるため、「全か無か」の判断を回避。定量的な根拠に基づく透明性の高い意思決定
- **短所:** 最適化モデル自体の前提（共分散推定等）に誤りがある場合、系統的な判断ミスにつながる

4.3.4 司会役による勝者判定（TradingAgents）

強気派と弱気派の議論が平行線に陥った場合、専門の\*\*議論ファシリテーター（Debate Facilitator）が介入する。ファシリテーターは議論履歴をレビューし、どちらの論理が優勢かを判定して支配的な視点（Prevailing Perspective）\*\*を決定・記録する。

- **長所:** 単なる多数決ではなく、論理の質に基づいた判定が可能
- **短所:** ファシリテーター自身の判定基準が暗黙的であり、再現性の検証が難しい

4.3.5 実績スコアに基づく選抜（FinRobot）

FinRobotの\*\*スマートスケジューラ（Smart Scheduler）\*\*は、過去のパフォーマンスが良いエージェントに優先的に決定権を付与する。ディレクターエージェントがタスクの内容を評価し、過去のパフォーマンス指標に基づいてエージェントをランク付けし、最もスコアが高い（適任である）エージェントにタスクを割り当てる。

- **長所:** 実績に基づく動的な権限委譲により、環境変化への適応力が高い
- **短所:** 過去のパフォーマンスが将来を保証しない（過剰適合リスク）

各方式の比較表

メカニズム	代表システム	決定速度	堅牢性	柔軟性	透明性
全会一致型	AlphaAgents	低	高	低	高
マネージャー決定型	TradingAgents / HedgeAgents	高	中	中	中
数理最適化型	HedgeAgents / MASFIN	中	高	高	高
司会役判定型	TradingAgents	中	中	中	低
実績スコア型	FinRobot	高	中	高	中



実際の運用においては、これらを単独で採用するのではなく、状況に応じて組み合わせるハイブリッド・アプローチが有効である。例えば、通常時は数理最適化型で迅速に判断し、市場急変時にはマネージャー決定型に切り替え、高確信度の銘柄については全会一致型で検証するといった階層的な設計が考えられる。

## 第5部: 5段階投資判断パイプラインの設計

第4部で定義したエージェントチームが「どのように協調するか」を設計したのに対し、本章ではそれらのエージェントが「どのような順序で何を判断するか」という投資判断パイプラインを設計する。MASFINの5段階Crewアプローチを基盤とし、AlphaAgentsの議論メカニズムを各段階に統合したモデルである。

フェーズ	役割	主要成果物
1. Postmortem	上場廃止・破綻銘柄の失敗パターン分析、生存者バイアス排除	リスク回避フラグ・失敗シグナル
2. Screening	市場全体から50-100銘柄を選別（アルファ・ファクトリー）	投資ユニバース・候補リスト
3. Analysis	3専門エージェントによる詳細分析、アルファ因子の特定	銘柄別スコア・深層レポート
4. Timing	エントリー時期の判定、先読みバイアスの排除	売買アクション（BUY/SELL/HOLD）
5. Portfolio	15-30銘柄のウェイト最適化、リスク調整後の最終配分	最終ポートフォリオ構成案

各段階の境界には\*\*要約エージェント（Summary Agent）\*\*が配置され、前段の分析結果を構造化して後段に引き渡す。これにより、情報の文脈を維持しつつ、HITL（Human-in-the-Loop）による検証ポイントが設けられる。

### 5.1 第1段階: ポストモルテム（過去の失敗からの学習）

AIの思考プロセスは、成功例ではなく「過去の失敗」から開始する。これは\*\*ポストモルテム（Failure Pattern Analysis）\*\*と呼ばれ、MASFINのPostmortem Crewが担当する。

#### 生存バイアスの回避

投資の世界において、現存する成功企業だけを分析対象にする「生存バイアス（Survivorship Bias）」は致命的な判断ミスを招く。成功企業のデータだけで学習したモデルは、破綻リスクの兆候を見逃す構造的欠陥を抱える。MASFINは、2020年から2025年にかけて倒産や上場廃止、コンプライアンス違反を犯した**18社**のデータをあえて分析に組み込む。

#### 失敗パターンの抽出

分析対象には、EVセクターのNKLA（Nikola）やRIDE（Lordstown Motors）、バイオテックのBLUE（Bluebird Bio）、小売のBBBYQ（Bed Bath & Beyond）などが含まれる。AIはこれらの「敗者」の記録から、以下の3つの警告シグナルカテゴリを抽出する。

失敗カテゴリ	具体例	検出指標
財務的不備	収益性の持続的欠如、上場維持基準の未達	営業利益率の長期マイナス推移、監査意見の限定
法令・規則違反	報告書未提出、コンプライアンス違反	SEC提出遅延、規制当局からの警告
センチメントの崩壊	ニュースの論調が急速にネガティブ化	センチメントスコアの急落、空売り比率の急増

これらの失敗シグナルは、後続のスクリーニング段階において「除外フラグ」として機能し、同様のパターンを示す銘柄を投資ユニバースから自動的に排除する。

5.2 第2段階: スクリーニング（アルファ・ファクトリー）

市場に存在する数千の銘柄を50-100程度に絞り込む段階では、AIは単なるフィルターではなく\*\*アルファ・ファクトリー（Alpha Factory）として機能する。11の主要な財務・研究文献からLLMを用いて抽出した100のシード・アルファ（Seed Alpha）\*\*が、このスクリーニングの基盤となる。

定性・定量の融合と信号抽出

スクリーニングでは、以下の2つのアルファ演算子を組み合わせてシグナルを生成する。

- **Cross-Section Operators** ( $f_{cs}$ ) : 同一時点における銘柄間の相対的關係（セクター内割安度等）を捉える
- **Time-Series Operators** ( $f_{ts}$ ) : 過去から現在に至るトレンドや時間的パターン（モメンタム等）を解析する

統合シグナルは以下の数式で表される：

$$\alpha_{ij}^{(t)} = w_{cs} \cdot f_{cs}(X_i^{(t)}) + w_{ts} \cdot f_{ts}(X_i^{(t)}, \dots, X_i^{(t-n)})$$

9つのアルファカテゴリ（Momentum、Mean Reversion、Volatility、Fundamental、Growth、Liquidity、Quality、Sentiment、Technical）から信号が抽出・フィルタリングされる。

選定プロセスの論理（3つの基準）

1. **アルファの適合性:** 現在の市場環境において、Information Coefficient（IC）が高いと期待されるファクターを保有しているか。バックテスト（SSE 50 Index）では、MASによる選択プロセスがMomentum因子の平均ICを0.0092から0.0208へと倍増させた
2. **リスクの事前排除:** 第1段階で学習した失敗パターンに合致する兆候がないか
3. **カテゴリーの多様性:** 特定のセクターやアルファカテゴリに偏らず、バランスの取れた候補群を形成しているか

5.3 第3段階: 詳細分析とアルファ・マイニング

候補銘柄が絞り込まれた後、第4部で定義した3つの専門エージェント（ファンダメンタル、センチメント、バリュエーション）が本格的な深掘り分析を実施する。

3つの専門エージェントの詳細分析

エージェント	分析対象と手法	使用ツール・数学的根拠
ファンダメンタル	10-K/10-Qレポートの精密解読。収益、キャッシュフロー、マージンの時系列分析	RAGツール、財務データ抽出API
センチメント	ニュース、アナリスト評価の定量化。反射型プロンプトによる批判的要約	Reflection-enhanced Summarization
バリュエーション	歴史的株価・出来高からの統計的指標算出（年率換算ボラティリティ等）	252取引日前提の年率換算計算ツール

RAG精度監視（Arize Phoenix）

Financial RAGツールの出力品質は、**Arize Phoenix**などの評価ツールで常時監視される。RAGが不適切なドキュメントを検索した場合や、コンテキストの欠落が検出された場合、エージェントに再検索を指示する自動フィードバックループが組み込まれている。

アルファ因子の特定

各エージェントの分析結果から、グロスマージンの改善トレンド、移動平均からの乖離度、センチメントスコアの変化率といったアルファ因子が抽出される。この段階でのエージェント間議論（第4部のラウンドロビンまたは弁証法的アプローチ）により、単一視点では見落とされるリスクや機会が浮かび上がる。

### 5.4 第4段階: タイミング判断（エントリー検証）

優れた銘柄を特定しても、「いつ買うか」の判断を誤れば損失を被る。第4段階は、先読みバイアス（Look-ahead Bias）を厳格に排除した上で、最適なエントリータイミングを判定する。

#### バイアス対策のメカニズム

MASFINは、**コンテンポラニアス・データ・スナップショット（Contemporaneous Data Snapshot）**と**固定された過去のヒストリカル・ウィンドウのみ**を使用するように設計されている。分析指標は特定の時点に固定されたデータのみアクセスを制限し、未来の情報が分析に混入するデータリークage（Data Leakage）を技術的に遮断する。

さらに、StockBenchの手法に基づく**\*\*汚染チェック（Contamination Check）\*\***を導入する。LLMのトレーニングカットオフ日以降のデータで検証を行い、モデルがその期間の株価推移を事前に「知っている」かどうかをテストする。正解率が高い期間は評価から除外することで、LLM固有のバイアスを排除する。

#### 先行指標の組み合わせ

AIは以下の4つの先行指標を組み合わせ、エントリーの妥当性を検証する。

指標	役割	判定基準
RSI-14	売られすぎ / 買われすぎのテクニカル判定	RSI < 30: 売られすぎ（買い候補）、RSI > 70: 買われすぎ（売り候補）
リターンZスコア	過去平均リターンからの統計的乖離度	$Z = \frac{R_t - \bar{R}}{\sigma_R}$ による異常リターンの検出
価格 vs 5日移動平均 (5DMA)	短期トレンドとの位置関係	MAからの乖離率による過熱・割安判定
出来高トレンド (Volume Trend)	取引の活発化による裏付け確認	出来高の増減と価格変動の整合性

これらの指標にSortino RatioやZ-Score、回帰スロープを組み合わせた総合スコアに基づき、BUY / SELL / HOLDのアクションが決定される。

### 5.5 第5段階: ポートフォリオ構築とリスク調整

最終段階では、前段までの分析結果を統合し、15-30銘柄のウェイトを最適化して最終ポートフォリオを構築する。

#### リスクプロファイルによる意思決定の違い

エージェントの挙動は、設定されたリスクプロファイルによって根本的に異なる。

リスクプロファイル	挙動	ポートフォリオ特性
Risk-Averse (リスク回避型)	パフォーマンスが良くてもボラティリティが高い銘柄はSELLを推奨。安定性と低ドロウダウンを最優先	低ベータ銘柄中心、ディフェンシブセクター重視

リスクプロファイル	挙動	ポートフォリオ特性
Risk-Neutral (リスク中立型)	ボラティリティよりもモメンタムを重視。高成長テック銘柄も積極的に組み入れ	高ベータ銘柄許容、グロースセクター混合

MASFINの検証によれば、Risk-Averse設定では高ボラティリティ銘柄が自動的に除外される一方、Risk-Neutral設定ではモメンタムを重視したポートフォリオが構築される。この動的な適応は、プロンプトエンジニアリングによるリスク許容度の指定を通じて実現される。

エージェント間議論による最終構成

最終ポートフォリオ案は、エージェント間の議論を経て決定される。この議論は単なる合意形成ではなく、ハルシネーションを抑制するためのセキュリティ機能として設計されている。

典型的な議論パターンは以下の通りである：

- **ファンダメンタル:** 「市場リーダーであり成長性は高い。買いた」
- **バリュエーション:** 「待て。5D MAからの乖離が大きく、ボラティリティも年率換算で基準を超えている」
- **合意形成:** 全員が納得するまで議論を繰り返し、バイアスを削ぎ落とした結論を導出

15-30銘柄への絞り込みとウェイト最適化

最終的なポートフォリオ構築では、以下の手法を組み合わせる。

**Dynamic Strategy Optimization:** 選定されたアルファ群のウェイト最適化に3層のMLP（多層パーセプトロン）を採用する。入力層にデイリー・アルファ値、中間層に10個のReLU活性化ノード、出力層に予測利回りを配置し、市場環境に応じてアルファファクターの重みを動的にシフトする。強気相場ではモメンタム系（RSI、MACD）、不確実性の高い相場ではボラティリティ系（ATR、Bollinger Bands）やファンダメンタル指標への自動調整が行われる。

ウェイト最適化の手法:

- **Top-k / Drop-nセレクション:** 全銘柄をアルファ値でランク付けし、上位k銘柄を採用。ターンオーバーコストを抑制するため、閾値を下回るまで保持
- **リスクパリティとの組み合わせ:** 等金額配分（Equal-weight）による分散効果と、ボラティリティ調整（Risk-parity）によるシャープレシオの向上を両立
- **週次リバランス:** MASFINプロトコルに基づき、週次でポートフォリオを再構築し、短期的な市場変動に対応

実績の検証

このパイプラインの有効性は、以下の実績データによって裏付けられる。

検証期間	手法	累積リターン	ベンチマーク
2023年1月 - 2024年1月（SSE 50）	Alpha Factory + Dynamic Strategy Optimization	53.17%	-11.73%（SSE 50指数）
8週間ライブ運用（MASFIN）	5段階パイプライン全体	7.33%	4.92%（S&P 500）

MASFINの8週間ライブ運用で記録した7.33%の累積リターンは、S&P 500（4.92%）およびNASDAQ-100を凌駕しており、マルチエージェント型5段階パイプラインの実効性を実証するものである。

第6部: アルファ生成とポートフォリオ最適化

## 6.1 アルファファクター自動生成（Seed Alpha Factory）

マルチエージェント・システム（MAS）が持続的なアルファを創出するうえで、その起点となるのが**Seed Alpha Factory**の構築である。従来のクオンツ運用では、アルファファクターの発見は研究者の仮説駆動型プロセスに依存しており、探索範囲が個人の知識と経験に制約されていた。本システムでは、LLM（大規模言語モデル）を活用し、11の主要な財務・研究文献からアルファファクター候補を自動的に抽出・生成するアプローチを採用する。

具体的には、LLMが学術論文や財務分析手法を解釈し、Pythonコードとして実行可能なアルファシグナルを自動生成する。このプロセスでは、まずLLMが文献中の投資仮説やアノマリーを構造化し、それをCross-Section（横断面）およびTime-Series（時系列）の演算子として定式化する。生成されたコードは即座にバックテスト環境で実行され、予測能力（Information Coefficient: IC）の統計的検証が行われる。ICが有意水準を超えないファクターは自動的に棄却され、残存したもののみがアルファ候補プールに登録される。

この手法により、HKUST（香港科技大学）の「Automate Strategy Finding」研究では**100のシードアルファ**が生成された。これらは網羅的にフィルタリング・選別され、最終的に高い予測精度を示すファクター群へと精製される。人間の認知バイアスや探索範囲の制約から解放されたこのプロセスは、アルファの「発見」を工業的なスケールで実現する点で画期的である。

## 6.2 9つのアルファカテゴリー

Seed Alpha Factoryから生成されたアルファは、以下の9つのカテゴリーに分類される。各カテゴリーは市場の異なる側面を捉え、多角的な予測シグナルを提供する。

カテゴリー	説明	LLM選択後の平均IC
Momentum	価格の継続的トレンドを捕捉。RSI、MACDなどに基づく	0.0208（ベースライン0.0092から倍増）
Mean Reversion	平均回帰傾向を利用。過度な乖離からの反発を予測	有意な改善を確認
Volatility	ボラティリティの構造変化を検出。ATR、Bollinger Bands など	有意な改善を確認
Fundamental	財務諸表から本源的価値を評価。利益率、キャッシュフロー	有意な改善を確認
Growth	売上高・利益の成長率トレンドを分析	有意な改善を確認
Liquidity	流動性リスクプレミアムを捕捉。出来高、売買回転率	有意な改善を確認
Quality	企業の財務健全性を評価。ROE、負債比率、利益の安定性	有意な改善を確認
Sentiment	ニュース・SNS・アナリスト評価から市場心理を定量化	有意な改善を確認
Technical	チャートパターンやテクニカル指標から売買シグナルを導出	有意な改善を確認

特筆すべきは、LLMによる選択プロセスを経ることで、各カテゴリーのICが劇的に向上する点である。例えばMomentum因子では、ベースラインの平均IC 0.0092がLLM選択後に0.0208へと倍増しており、VolatilityやFundamentalにおいても同様の改善が確認されている。これは、LLMが単なるランダム選択ではなく、市場の構造的パターンを理解したうえでファクターを選別していることを示唆する。9カテゴリーの均衡を保つことは、特定の市場レジームへの過度な依存を回避し、ポートフォリオ全体のロバスト性を担保する上でも重要である。

## 6.3 動的ウェイト最適化（3層MLP）

選定されたアルファファクター群を最終的なポートフォリオ・ウェイトに変換するプロセスには、\*\*3層MLP（多層パーセプトロン: Multi-Layer Perceptron）\*\*が採用される。このニューラルネットワークは、市場環境の変化に応じてアルファファクターの重みを動的に調整する「ウェイト・ゲーティング（Weight Gating）」の役割を果たす。

### アーキテクチャの構成:

- **入力層:** 各アルファファクターのデیلیー値（9カテゴリーからのシグナル）
- **中間層（隠れ層）:** 10個のReLU（Rectified Linear Unit）活性化ノード
- **出力層:** 予測利回り（各銘柄の期待リターン）

このアーキテクチャの核心は、市場環境に応じた**動的ウェイトシフト**にある。強気相場（Bull Market）ではMomentum系ファクター（RSI、MACDなど）への重みが増大し、不確実性の高い相場（High Uncertainty）ではVolatility系（ATR、Bollinger Bands）やFundamental指標への重みが自動的にシフトする。これにより、固定ウェイトのモデルでは対応困難な市場レジームの転換（Market Regime Shift）に柔軟に適応する。

アルファシグナルの統合は、以下の2つの演算子を基盤とする:

1. **Cross-Section演算子 ( $f_{cs}$ ):** 同一時点における銘柄間の相対的關係（セクター内割安度、相対的モメンタムなど）を捉える
2. **Time-Series演算子 ( $f_{ts}$ ):** 過去から現在に至るトレンドや時間的パターン（移動平均、ボラティリティクラスタリングなど）を解析する

これらを統合した最終的なアルファシグナルは、以下の数式で表される:

$$\alpha_{ij}^{(t)} = w_{cs} \cdot f_{cs}(X_i^{(t)}) + w_{ts} \cdot f_{ts}(X_i^{(t)}, \dots, X_i^{(t-n)})$$

ここで、 $w_{cs}$  と  $w_{ts}$  はそれぞれCross-SectionおよびTime-Series演算子の重みであり、MLPが市場環境に応じて動的に決定する。HKUSTの研究では、この手法によりSSE 50指数において**53.17%の累積リターン**（2023年1月-2024年1月）を記録し、同期間のベンチマーク（-11.73%）を圧倒的に凌駕する成果を達成している。

## 6.4 HedgeAgentsの予算配分最適化

HedgeAgentsフレームワークでは、マネージャーAI（Otto）が各専門エージェント（ビットコイン担当、株式担当、FX担当など）からの分析報告を集約し、数理最適化モデルを用いて最適な投資比率を決定する「予算配分会議（Budget Allocation Conference）」を実施する。

### ステップ1: 期待リターンの推定

各エージェントの報告書（ $R$ ）と補助ツール（ $R_a$ ）を基に、LLMが各資産  $i$  の将来の期待リターン（ $\rho_i$ ）を予測する。この段階で、Fundamental Agent、Sentiment Agent、Valuation Agentそれぞれの分析結果が統合される。

### ステップ2: 指標の計算

以下の3つの指標を算出する:

- **期待トータルリターン ( $I_{etr}$ ):** 各資産の期待リターンと重みの加重平均

$$I_{etr} = \sum \omega_i \rho_i$$

- **ポートフォリオ全体のリスク ( $I_{pr}$ ):** 資産間の共分散  $\sigma_{ij}$  を考慮した統合リスク

$$I_{pr}^2 = \sum \sum \omega_i \omega_j \sigma_{ij}$$

- **条件付きドローダウンリスク ( $I_{cvar}$ ):** 極端な市場環境における尾部リスク（Conditional Value at Risk）

### ステップ3: 最適化目的関数

リスクを抑制しつつリターンを最大化する最適ウェイト（ $\omega^*$ ）を以下の目的関数により決定する:

$$\omega^* = \arg \max_{\omega} (I_{etr} - \lambda_1 I_{pr} - \lambda_2 I_{cvar})$$

ここで  $\lambda_1, \lambda_2$  はリスク回避係数であり、投資家のリスク許容度を反映するパラメータである。制約条件として、予算の合計が 100% ( $\sum \omega_i = 1$ )、各配分が非負 ( $\omega_i \geq 0$ 、空売り禁止) を課す。この定式化により、期待リターンの追求と下方リスクの制御を同時に実現する。

6.5 FinRobotの強化学習アプローチ

FinRobotフレームワークでは、HedgeAgentsの静的最適化とは対照的に、\*\*マルコフ決定過程 (MDP: Markov Decision Process) \*\*に基づく強化学習 (FinRL) を採用し、動的にポートフォリオ配分を最適化する。

MDPの構成要素は以下の通りである:

- **状態 ( $s_t$ ):** 現在の市場環境 (株価、ボラティリティ、マクロ指標など)
- **行動 ( $a_t$ ):** 売買判断や配分変更
- **報酬 ( $r$ ):** 行動の結果として得られる収益
- **遷移:** 行動により次の状態  $s_{t+1}$  へ移行

目的関数は、将来にわたって得られる累積報酬の期待値を最大化する方策 (Policy  $\pi_\theta$ ) の学習である:

$$J(\pi_\theta) = \mathbb{E} \left[ \sum \gamma^t r(s_t, a_t, s_{t+1}) \right]$$

ここで  $\gamma$  は割引率であり、将来の報酬を現在価値に換算するパラメータである。

動的最適化 vs. 静的最適化の比較:

項目	HedgeAgents (静的最適化)	FinRobot (動的最適化)
手法	数理最適化 (凸最適化)	強化学習 (Deep RL)
配分決定	一時点の最適解を算出	市場との相互作用を通じて逐次学習
市場適応	リバランス時に再計算	リアルタイムで方策を更新
強み	解釈性が高く、制約条件の明示が容易	非線形な市場ダイナミクスへの適応力
弱み	市場レジーム変化への反応が遅い	学習の不安定性、ブラックボックス化のリスク

実務上は、HedgeAgentsの静的最適化を基盤としつつ、FinRobotの動的最適化を補完的に活用するハイブリッドアプローチが有効である。これにより、解釈可能性を維持しながらも、市場環境の急変に対する適応力を確保できる。

第7部: バイアス排除とバックテスト設計

7.1 先読みバイアス (Look-ahead Bias) の排除

LLMベースのMAS運用において最も致命的なバイアスは、\*\*先読みバイアス (Look-ahead Bias) \*\*である。これは、エージェントが意思決定時点では利用不可能であった将来の情報を推論の前提としてしまうことで発生する。従来のクオンツ手法における単純な時系列の取り違えとは異なり、MASでは非構造化データの自律的な取得・解釈を行うため、バイアスの混入がはるかに不可視かつ巧妙に発生する。

時間的隔離 (Temporal Isolation) の実装:

各エージェントのデータアクセス権限を時間軸で厳格に制限する。Fundamental Agentは検証対象日以前に受理 (Filed) された 10-K/10-Qのみを参照し、Sentiment Agentは検証対象日の市場クローズ (16:00 EST等) 以前に発行されたニュースのみにアクセスを許可される。Valuation Agentは過去の固定ウィンドウ (例: 直近21日間) のOHLCVデータに厳格に制限される。

デバイアス処理 (Debiasing) :

エージェントのプロンプトには「あなたは〇〇年〇月〇日にいる」という時間軸のペルソナを設定し、将来の知識を推論に使用することを明示的に禁止するメタインストラクションを付与する。これにより、エージェントの出力から結果論的な表現（「その後〇〇が起きた」等）を排除し、中立表現への書き換えを強制する。

#### MASにおけるバイアス混入の3つの具体的なシナリオ:

1. **学習データのカットオフによる未来知識:** GPT-4oなどのLLMは広範な学習データを持つ。2023年のバックテスト実行時に、モデルが2024年の企業倒産情報を既に保持している場合、その知識が「リスク分析」として暗黙的に推論に漏洩する
2. **RAGを通じた非決定的な情報取得:** タイムスタンプ制御のないRAGツールが、検証対象日より後に発行されたアナリストレポートを検索し、その「下方修正」情報をエージェントに提供してしまうケース
3. **市場反応後のニュース要約:** 株価急落の「理由」を解説する数日後のニュース記事を、Sentiment Agentが「急落の予兆」として読み込み、推論を自己正当化するケース

## 7.2 生存者バイアス（Survivorship Bias）の排除

上場廃止企業を含めないバックテストは、リターンを体系的に過大評価する。MASFINの「**Postmortem Crew**（事後検証班）」は、この生存者バイアスを排除するために設計された専門チームである。

#### Postmortem Crewの役割:

破綻・上場廃止に至った企業の「失敗パターン」を体系的に分析し、その知見をスクリーニングおよびポートフォリオ構築の議論の前提として組み込む。これにより、成功企業のみ偏った学習（Survivorship Bias）を構造的に防止する。

#### 具体的な対象企業と教訓:

- **NKLA (Nikola Corporation)** : EV関連のハイプサイクルにおける過大評価と、その後の不正会計発覚に至るプロセスの分析
- **BBBYQ (Bed Bath & Beyond)** : 小売業界の構造変化に適応できなかった経営戦略の失敗パターン
- **REV (Revlon)** : 過剰債務と競争激化による財務悪化から倒産に至るプロセスの分析

#### ビジネス上のインパクト:

失敗企業のパターンをエージェントが認識することは、単なるバックテスト精度の向上に留まらない。ドローダウンの深さを現実的な範囲に収めることで、**GP/LP間の信頼関係を維持し、パニックによる資産流出（ファンド償還）を未然に防ぐ**という、運用継続性に直結するリスク管理機能（Downside Protection）を果たす。

## 7.3 データ汚染（Contamination）への対応

LLMの学習データに含まれる過去の市場情報が、バックテスト結果を汚染するリスクへの対応は、MAS運用固有の課題である。

#### StockBench手法: LLMの株価言い当てテスト

StockBenchは、LLMが特定の過去の株価データを「記憶」しているかどうかを検証するベンチマークである。LLMに対して特定日時の株価を直接質問し、その回答精度を測定する。モデルのトレーニングカットオフ日以前のデータについて高い回答精度を示す場合、そのモデルは当該期間のバックテストには不適切であると判断される。

#### トレーニングカットオフ日以降のデータ検証:

バックテスト対象期間は、使用するLLMのトレーニングカットオフ日以降に設定することが原則である。カットオフ日以前の期間をバックテスト対象とする場合は、StockBench等の手法でデータ汚染の程度を定量的に評価し、結果の解釈に反映させる必要がある。

## 7.4 ポイント・イン・タイム（PiT）データ管理

バックテストの妥当性は、各分析ステップにおけるデータの時間的整合性の維持に依存する。



同時性のスナップショット (Contemporaneous Snapshots) :

すべての分析指標は、特定の時点に固定されたデータのみアクセスを制限する。財務データのタイムスタンプは市場時間 (Open/Close) と秒単位で整合させ、Data Leakage (情報漏洩) を技術的に遮断する。例えば、決算発表が市場クローズ後に行われた場合、その情報は翌営業日のスナップショットにのみ反映される。

時系列セグメンテーション (TS/VS/Test) :

訓練 (Training Set)、検証 (Validation Set)、テスト (Test Set) の各フェーズを時間軸で完全に分離する。フェーズ間の重複を一切許容しないことで、モデルが将来のボラティリティ構造や市場レジームを学習することを防止する。「Automate Strategy Finding」の手法に倣い、以下のような時間分割を実施する:

- **Training Set (TS):** モデル学習期間 (例: 2018年1月 - 2021年12月)
- **Validation Set (VS):** ハイパーパラメータ調整期間 (例: 2022年1月 - 2022年12月)
- **Test Set:** アウトオブサンプル評価期間 (例: 2023年1月 - 2024年1月)

7.5 RAG/ベクトルDBの時間フィルタリング

RAG (Retrieval-Augmented Generation) ツールにおいては、検索クエリのAPIレイヤーおよびベクトルデータベース (Arize Phoenix等) において、メタデータフィルタリングを強制しなければならない。エージェントが「今がいつか」を認識するだけでは不十分であり、物理的に未来の文書に触れられない環境を構築することが不可欠である。

コード例 (概念図) :

```
# ベクトルDB検索時にメタデータフィルタを強制
metadata_filter = {
    "timestamp": {"$lte": context.backtest_date},
    "document_status": "official_release"
}
results = vector_db.search(query, filter=metadata_filter)
```

加えて、Arize Phoenix等のオブザーバビリティツールを用いて、RAGの「Faithfulness (誠実性)」と「Relevance (関連性)」を常時監視する。要約ツールが将来のイベント (決算発表後の解説記事など) を誤って抽出していないかをリアルタイムに検知し、違反が検出された場合はエージェントの推論を即座に無効化する。RAGが参照した全ドキュメントのソース名と発行日時をログに保存することで、事後的な監査トレースも確保する。

7.6 堅牢な評価指標

累積リターンは最終結果に過ぎない。戦略の頑健性を多面的に評価し、バイアスの有無を逆検知するために、以下の指標群を体系的に運用する。

指標	定義	目的	戦略的閾値
Information Coefficient (IC)	Corr(予測, 実績)	アルファシグナルの純粋な予測精度の測定	> 0.02 (LLM選択後)
Sharpe Ratio	$(R_p - R_f) / \sigma_p$	リスク単位あたりの超過収益の評価	> 0.28 (SSE 50バックテスト)
Sortino Ratio	$(R_p - R_f) / \sigma_{downside}$	下方リスクのみを考慮したリスク調整後リターン	> 0.20
Calmar Ratio	$R_{ann} / MDD$	年率リターンと最大ドローダウンの比率	> 1.05

指標	定義	目的	戦略的閾値
Maximum Drawdown (MDD)	ピークからの最大下落率	テールリスクの評価	戦略・リスク許容度に依存
Annualized Return	$((1 + R_{cum})^{252/n}) - 1$	長期的なアルファ創出能力の絶対値	ベンチマーク超過
Annualized Volatility	$\sigma_{daily} \times \sqrt{252}$	エージェント判断の不確実性とリスク量	目標リスク水準以下

Rolling Sharpe Ratio:

4ヶ月から1年のローリング・ウィンドウでSharpe Ratioを計測し、特定期間の「ラッキーパンチ」を排除する。時系列的に安定した超過収益を達成しているか、あるいは特定の市場環境でのみ有効な戦略であるかを識別する。

市場レジーム分析:

リスク許容度設定 (Risk-neutral vs. Risk-averse) ごとのパフォーマンス乖離を分析する。AlphaAgentsのH1 2024における実験では、**Risk-averse (リスク回避型) エージェントが上昇相場においてボラティリティの高いテクノロジー銘柄を除外し、ベンチマークを下回った**。この結果を「失敗」ではなく「意図したリスク特性の反映」と解釈できる透明性が、MAS運用には不可欠である。異常に高いICがデータリークに起因するものではないか、感度分析を実施することも重要なバイアス検出手法である。

7.7 実装チェックリスト

先読みバイアス・ゼロを目指すための非妥協的な運用プロトコルを以下に示す。

データ・エンジニアリング (Data Engineering)

- ☐ バックテスト全期間でポイント・イン・タイム (PiT) データが整備されているか
- ☐ 上場廃止企業 (NKLA、BBBYQ、REV等) がユニバースに含まれているか
- ☐ 訓練・検証・テストの各期間が時系列で厳格に分離されているか (TS/VS/Testの重複禁止)
- ☐ 財務・ニュースデータのタイムスタンプが市場時間 (Open/Close) と秒単位で整合しているか

エージェント・プロンプト設計 (Agent Prompt Design)

- ☐ 「あなたは〇〇年〇月〇日にいる」という時間軸のペルソナが設定されているか
- ☐ 将来の知識を推論に使用することを明示的に禁止するメタインストラクションがあるか
- ☐ 各エージェント (Fundamental、Sentiment、Valuation、Risk Adjuster) の役割と参照可能データ範囲が明示されているか
- ☐ リスク許容度設定 (Risk-averse / Risk-neutral) がポートフォリオの目的と整合しているか

RAG/ツール呼び出し (RAG/Tool Usage)

- ☐ 検索時に、検証対象日以降の文書を除外するメタデータフィルタが強制されているか
- ☐ Arize Phoenix等を用いて「Faithfulness (誠実性)」と「Relevance (関連性)」を常時監視しているか
- ☐ 要約ツールが将来のイベント (決算発表後の解説記事等) を抽出していないか
- ☐ RAGが参照した全ドキュメントのソース名と発行日時がログに保存されているか

事後検証プロトコル (Post-test Validation)

- ☐ Human-in-the-Loop (HITL) で、エージェントが後知恵的な推論をしていないかレビューしたか
- ☐ Global Mean Benchmarking (MASFIN方式) を用い、コホート平均に対する相対的な優位性を確認したか
- ☐ 異常に高いICがデータリークに起因するものではないか、感度分析を行ったか
- ☐ 特定の市場レジーム (例: H1 2024の上昇相場) におけるリスク設定ごとのパフォーマンス乖離を論理的に説明できるか

- ☐ StockBench等の手法でLLMのデータ汚染（Contamination）の有無を検証したか
- ☐ 生存者バイアス排除後のダウンロードが、GP/LP間の信頼関係を維持できる範囲に収まっているか
- ☐ 討論（Debate）ログが監査トレース（Audit Trace）として保存され、全投資判断の根拠が追跡可能か

## 第8部: 本プロジェクトへの統合設計

本部では、学術研究の知見と本プロジェクト固有の文脈——すなわちファンドマネージャー Y の投資哲学、既存の finance パッケージ群、Claude Code によるエージェントオーケストレーション基盤——を接続し、4段階の統合ロードマップを設計する。

### 8.1 Stage 1 成果の活用 --- Y の判断軸と Critique Agent

Stage 1（完了済み）では、Phase 2 の5銘柄（CHD, COST, LLY, MNST, ORLY）に対する Y の確信度スコアリングと feedback.md の定性フィードバックから、2つの中核ドキュメントが生成された。

- **judgment\_patterns.md**: 5銘柄横断の判断パターン分析。却下パターン7類型（結果と原因の取り違い、一般論、差別化不明、事実誤認、飛躍的解釈、定量裏付けなし、戦略と優位性の混同）と高評価パターン6類型（定量的裏付け、直接的 CAGR 接続メカニズム、能力 > 結果、明確な因果チェーン、他社事例との類推、構造的 vs 補完的の区別）を体系化。
- **dogma\_draft.md**: K の Competitive\_Advantage\_Dogma.md のフォーマットを参考に、Y 固有の投資哲学を文書化したドラフト。

**Critique Agent への組み込み設計**: Y の却下パターン7類型は、ルールベースのチェックリストとして Critique Agent のシステムプロンプトに直接埋め込む。具体的には、AI が生成した競争優位性の主張に対し、以下の逐次検証を強制する。

1. **因果方向チェック**: 「この主張は結果を原因として提示していないか？」（例: 高シェアは結果であり、背後要因が示されていないか）
2. **一般性チェック**: 「この主張は同業他社にも等しく当てはまらないか？」（例: 「メガファーマなら誰でも持っている」パイプラインは優位性ではない）
3. **定量性チェック**: 「この主張には具体的な数値裏付けがあるか？」（例: 退職率9% vs 業界20%+ のような比較データの有無）
4. **CAGR 接続チェック**: 「この優位性から CAGR への接続メカニズムは明示されているか？」

Phase 2 の5銘柄データは検証セット（Validation Set）として活用する。Critique Agent が生成するスコアと Y の実際のスコアとの相関係数を測定し、0.7以上を初期目標とする。スコア乖離が大きい項目については、Y の詳細コメントと照合して判断ルールの精緻化を行う。

### 8.2 Stage 2 設計: Y 視点 Critique Agent の構築

Stage 2 の目標は、Y の暗黙知をシステムプロンプトとして形式化し、AI が Y の視点で投資判断を批評できる単一エージェントを構築することである。

**システムプロンプト設計**: Y の判断軸を以下の構造でプロンプトに組み込む。

#### [ペルソナ定義]

あなたはファンドマネージャー Y の投資判断基準を内面化した批評エージェントです。

#### [スコアリング体系]

5段階の確信度（90%: 構造的優位性+明確なCAGR接続+定量的裏付け ~ 10%: 事実誤認・競争優位性として不成立）

#### [却下ルール（Hard Rules）]

– 結果を原因と取り違えている主張 → 30%以下

- どの企業にも言える一般論 → 30%以下
- 戦略を優位性と混同 → 30%以下
- 事実誤認を含む → 10%

[高評価条件 (Soft Rules) ]

- 定量的裏付けあり → +20%
- 直接的CAGR接続メカニズムが明示 → +20%
- 能力（プロセス）に着目し結果ではない → +10%

**精度測定:** Phase 2 の全スコアリングデータ（5銘柄 x 各5-8項目 = 約30データポイント）に対し、Critique Agent のスコアと Y のスコアとのピアソン相関係数およびスコア差の平均絶対誤差（MAE）を計測する。MAE が15ポイント以内（例: Y が70% のとき Agent が55-85%の範囲）であれば実用水準と判断する。

**批判自動生成の例:** COST の Phase 2 データに対し、Critique Agent は「会員制モデルによる顧客ロックインは多くのサブスクリプション企業が採用しており、Costco 固有の構造的優位性として成立するためには、90%を超える更新率の背後要因（例: 年会費に対する perceived value の定量比較）が必要」といった Y スタイルの批判を自動生成する。

**アナリストごとのフレームワーク切替:** 各アナリストの判断軸ファイルを `analyst/Competitive_Advantage/analyst_{name}/dogma.md` に配置し、システムプロンプトをモジュール化する。K モードでは `Competitive_Advantage_Dogma.md` の7 Powers フレームワークを適用し、Y モードでは `dogma_draft.md` の因果チェーン重視型フレームワークを適用する。ファンドマネージャーが分析時にモードを指定することで、異なる視点からの批評を並列取得できる。

### 8.3 Stage 3 設計: 弁証法的マルチエージェントアーキテクチャ

Stage 3 では、Stage 2 の単一 Critique Agent を7つの専門エージェントに分解し、TradingAgents の Bull/Bear 弁証法アーキテクチャと AlphaAgents の専門エージェント間議論（Debate）メカニズムを統合する。

**エージェント構成と役割定義:**

エージェント	役割	学術的対応
FM Agent (Y's Philosophy)	全体統括。最終的な投資判断を Y の哲学に基づき下す	AlphaAgents の Portfolio Manager
Bull Agent	買いの根拠を徹底探索。楽観シナリオの論理構築	TradingAgents の Bull Researcher
Bear Agent	売りの根拠を徹底探索。リスクシナリオの具体化	TradingAgents の Bear Researcher
Competitive Advantage Critic	Y の判断軸に基づき優位性の真贋を批判	AlphaAgents の Fundamental Agent を Y 哲学で拡張
CAGR Validator	数値の根拠と銘柄間の一貫性を検証	MASFIN の Analysis Crew の定量検証機能
Causality Checker	結果 vs 原因の区別を強制。因果方向の論理検証	本プロジェクト固有（Y の課題5への対応）
Historical Analogy Agent	Y の過去判断事例との整合性チェック	HedgeAgents の Historical Pattern Module に相当
Bias Detector	確認バイアス・初期判断への固執を検出	TradingAgents の Risk Management Team に相当

**議論プロトコル:** AlphaAgents の Round Robin 方式を採用する。各銘柄分析において、以下の3ラウンド構成で議論を実行する。

- **Round 1 (独立分析)** : Bull Agent と Bear Agent が独立にレポートを生成。Competitive Advantage Critic が Y の判断軸に基づく初期批評を付与。
- **Round 2 (相互批評)** : 各エージェントが他エージェントのレポートを読み込み、反論または支持を表明。Causality Checker と CAGR Validator が横断的な整合性チェックを実行。Historical Analogy Agent が Y の過去判断事例 (Phase 2 の5銘柄における具体的判断) との類似性を指摘。
- **Round 3 (合意形成)** : FM Agent が全議論ログを統合し、Y のスコアリング体系に基づいて最終判断を下す。Bias Detector が議論全体を通じた認知バイアスの有無を報告。

この設計により、Zscaler のケーススタディで実証された「当初 BUY を提示したエージェントが他者の指摘を受けて SELL に転じる」プロセスが、Y の投資哲学という固有のフィルターを通して再現される。

## 8.4 Stage 4 設計: 完全 AI 運用チーム

Stage 4 は、MASFIN の5段階パイプライン (Postmortem → Screening → Analysis → Timing → Portfolio) を完全実装し、リサーチから銘柄選定、ポートフォリオ構築までを自動化する最終段階である。

### MASFIN 統合パイプラインの実装:

1. **Postmortem Crew:** 上場廃止銘柄 (NKLA, RIDE, BBBYQ 等) の失敗シグナルをデータベース化し、生存者バイアスを排除。edgar パッケージを用いて過去の SEC Filings から破綻前兆パターンを抽出。
2. **Screening Crew:** market パッケージの yfinance 連携と FRED 連携を活用し、マクロ経済指標と個別銘柄の財務状況に基づき、投資ユニバース (50-100銘柄) を自動選別。
3. **Analysis Crew:** Stage 3 の弁証法的マルチエージェントが銘柄別の深層分析を実行。analyze パッケージのテクニカル分析機能と統合。
4. **Timing Crew:** Sortino Ratio、Z-Score、回帰スロープに基づく売買タイミング判定。market パッケージのリアルタイムデータ取得機能と連携。
5. **Portfolio Crew:** factor パッケージのマルチファクターモデルと strategy パッケージのリスク計算・ポートフォリオ管理機能を統合。Black-Litterman モデルにエージェントの確信度スコアを反映。

**リアルタイム市場データ連携:** yfinance API および Bloomberg 連携による日次データ取得を基盤とし、rss パッケージのフィード監視機能でニュースイベントをトリガーとした臨時分析を実行する。

## 8.5 既存 finance パッケージとの統合

本プロジェクトの既存パッケージ群は、マルチエージェントシステムのツールセットとして以下のように統合される。

パッケージ	統合先エージェント	提供機能
market	Bull/Bear Agent, Timing Crew	yfinance による株価・為替・指数データ取得、FRED による経済指標取得、Bloomberg 連携、キャッシュ機能
analyze	CAGR Validator, Analysis Crew	テクニカル分析 (移動平均、RSI、MACD、ボリンジャーバンド)、統計分析、セクター分析、可視化
factor	Portfolio Crew	バリュー/モメンタム/クオリティ/サイズ/マクロの5ファクターモデル。HKUST の Seed Alpha 自動生成と対応する因子構成
strategy	Portfolio Crew, FM Agent	リスク計算 (VaR, CVaR)、ポートフォリオ最適化、リバランス分析

パッケージ	統合先エージェント	提供機能
edgar	Fundamental Agent, Postmortem Crew	SEC Filings (10-K, 10-Q, 8-K) のテキスト抽出・セクション分析。 AlphaAgents の Financial RAG Tool に相当
rss	Sentiment Agent, Screening Crew	ニュースフィード監視、記事抽出、キーワード検索
news	Sentiment Agent	ニュース処理パイプライン、フィルタリング、GitHub Issue 連携
database	全エージェント共通	SQLite/DuckDB 接続、構造化ロギング、日付ユーティリティ

この統合により、AlphaAgents が定義する3エージェント（Fundamental, Sentiment, Valuation）のツールセットを、既に実装・テスト済みの Python パッケージで実現できる。新規開発が必要な部分は、エージェント間の議論プロトコルとオーケストレーション層に限定される。

## 8.6 Human-in-the-Loop（HITL）設計

完全自動化の罟を回避し、規制要件を満たすため、MASFIN の HITL モデルを本プロジェクトに適合させる。

**Summary Agent による検証ポイント:** 各フェーズの境界に Summary Agent を配置し、以下の3つの検証ポイントを設ける。

- 1. **Screening 後:** 投資ユニバースの妥当性確認。セクター偏りやサイズ偏りの検出。
- 2. **Analysis 後:** 弁証法的議論の結論と、Y の過去判断との整合性確認。ファンドマネージャーが議論ログを閲覧し、エージェントの推論品質を評価。
- 3. **Portfolio 構築後:** 最終ポートフォリオ案のリスク特性確認。集中度、セクターエクスポージャー、ファクターティルトの検証。

**議論ログの監査トレース（Audit Trace）:** エージェント間の全議論ログを JSON 形式で永続化し、投資判断の根拠を遡及可能にする。これは金融規制における説明責任（Explainability）要件への対応であると同時に、Y の判断軸の継続的改善のためのフィードバックデータとなる。

**計算コスト最適化:** Claude Code の API コールを最小化するため、以下の戦略を採用する。(1) 定型的な数値計算（テクニカル指標、財務比率）は Python パッケージ内で完結させ、LLM は定性的判断にのみ使用。(2) 議論ラウンド数を固定（3ラウンド）し、合意に達した場合は早期終了。(3) 異常値（スコア乖離が30%以上）が発生した場合のみ追加ラウンドを実行。

# 第9部: 実装ロードマップ

## 9.1 短期計画（Phase 1: 3ヶ月）

Phase 1 の目標は、Critique Agent プロトタイプの構築と Y の判断軸の検証である。

**マイルストーン:**

- **Month 1:** Y の判断軸 (dogma\_draft.md) をシステムプロンプトに組み込んだ Critique Agent の初版を Claude Code のエージェント定義 ([.claude/agents/](#)) として実装。Phase 2 の5銘柄データを入力し、Y のスコアとの相関を測定。
- **Month 2:** Few-shot Examples の再設計。Y が高評価した分析を「良い例」、却下した分析を「悪い例」として選定し、プロンプトに組み込む。CAGR パラメータ処理のルール統一（優位性確信度と CAGR 接続は独立評価）を実装。

- **Month 3:** Critique Agent の精度改善イテレーション。Y からのフィードバック (O/x/Δ) を反映し、判断ルールを精緻化。目標: ピアソン相関係数 0.7 以上、MAE 15ポイント以内。

**技術要件:** Claude Code エージェント定義、既存の analyst/ ディレクトリのデータ、Phase 2 スコアリングデータ。新規パッケージ開発は不要。

## 9.2 中期計画 (Phase 2: 6ヶ月)

Phase 2 の目標は、弁証法的マルチエージェントシステムの構築と既存パッケージ統合である。

**マイルストーン:**

- **Month 4-5:** 7エージェント (FM Agent, Bull Agent, Bear Agent, Competitive Advantage Critic, CAGR Validator, Causality Checker, Historical Analogy Agent, Bias Detector) の個別実装。各エージェントのシステムプロンプトと役割定義。Claude Code の Agent Teams 機能を活用した並列実行基盤の構築。
- **Month 5-6:** Round Robin 議論プロトコルの実装。3ラウンド構成の議論フローをスキル定義 (`.claude/skills/`) として実装。議論ログの JSON 永続化機能。
- **Month 6-7:** market パッケージ (yfinance, FRED)、analyze パッケージ (テクニカル分析)、edgar パッケージ (SEC Filings) とのツール統合。各エージェントが Python パッケージの関数を直接呼び出せるインターフェースの構築。
- **Month 7-9:** バックテストエンジンの構築。strategy パッケージのリスク計算機能を活用し、過去データに対するエージェントの投資判断のパフォーマンスを測定。Temporal Alignment (時間的整合性) を厳格に実装し、先読みバイアスを排除。

## 9.3 長期計画 (Phase 3: 12ヶ月)

Phase 3 の目標は、フルポートフォリオ最適化と自律運用の実現である。

**マイルストーン:**

- **Month 10-12:** MASFIN 5段階パイプラインの完全実装。Postmortem Crew による生存者バイアス排除、Screening Crew による投資ユニバース自動選別。
- **Month 13-15:** 3層 MLP による Dynamic Strategy Optimization の実装。factor パッケージの5ファクターモデルを入力層とし、市場レジームに応じたウェイト動的制御を実現。
- **Month 16-18:** Black-Litterman モデルの統合。エージェントの確信度スコアを「ビュー (Views)」として共分散行列に反映し、平均分散最適化を自律実行。
- **Month 19-21:** リアルタイム運用とモニタリングシステムの構築。rss パッケージによるニュースイベント駆動の臨時分析、日次ポートフォリオリバランスの自動化。
- **継続的:** 自己反省 (Reflection) ループの実装。エージェントの過去判断とその後の実績を照合し、システムプロンプトとルールベースの継続的改善を自動化。

## 9.4 技術スタック

レイヤー	技術	用途
エージェントオーケストレーション	Claude Code (Agent Teams, Skills, Commands)	マルチエージェント間の議論制御・タスク分配
言語・ランタイム	Python 3.12+, uv	パッケージ管理・依存関係解決
市場データ	yfinance API, FRED API, Bloomberg 連携	株価・経済指標・ニュースデータ取得

レイヤー	技術	用途
SEC Filings	edgartools (edgar パッケージ経由)	10-K/10-Q/8-K のテキスト抽出・分析
知識管理	RAG / ベクトル DB (将来導入)	Y の過去判断事例の検索・類推
既存パッケージ	market, analyze, factor, strategy, edgar, rss, news, database	前述の統合設計に基づく各エージェントのツールセット
品質管理	Ruff, pyright, pytest + Hypothesis	コード品質・型安全性・テスト

## 第10部: リスクと課題

### 10.1 技術的リスク

**LLM のハルシネーション:** マルチエージェント議論はハルシネーション抑制に有効であることが AlphaAgents と TradingAgents の双方で実証されているが、完全な排除は不可能である。特に、財務数値の引用において誤った数値を生成するリスクが残る。対策として、CAGR Validator による数値の外部データソース照合と、定型的な計算は Python パッケージ内で完結させる設計を採用する。

**プロンプトの脆弱性:** Y の判断軸をシステムプロンプトに組み込む設計は、プロンプトの微細な表現変更がエージェントの判断に大きく影響する「プロンプト感度 (Prompt Sensitivity)」問題を内包する。Phase 2 データでの継続的な回帰テストと、Few-shot Examples の影響度測定 (MSFT/TSMC 事例で既に確認済み) による管理が必要である。

**API 依存とレートリミット:** yfinance API、FRED API、Claude API への依存は、レートリミットやサービス障害時のシステム停止リスクを伴う。既存の market パッケージに実装済みのキャッシュ機能とポライトディレイ (Polite Delay)、RateLimitError 変換を活用し、グレースフルデグラデーション (Graceful Degradation) を設計する。

**モデルバージョン変更時の一貫性:** Claude のモデル更新により、同一プロンプトに対する出力の一貫性が損なわれるリスクがある。Phase 2 データに対する定期的な回帰テストを自動化し、スコア相関の低下を早期検出する。

### 10.2 運用リスク

**バックテストと実運用のギャップ:** HKUST の研究が SSE 50 において53.17%の累積リターンを記録した一方、StockBench の検証ではリアル市場でのパフォーマンスがバックテスト結果を下回ることが報告されている。先読みバイアスの排除 (Temporal Alignment) と、MASFIN プロトコルに準じたライブ検証期間の設定が不可欠である。

**市場レジーム変化への適応遅延:** 急激な市場環境の変化 (金融危機、地政学リスクの顕在化) に対し、エージェントの学習済みパターンが陳腐化するリスクがある。Dynamic Strategy Optimization (3層 MLP) による動的ウェイト調整と、Postmortem Crew による新たな失敗パターンの継続的学習で対応する。

**データ品質と遅延:** yfinance 等の無料データソースは、機関投資家向けデータプロバイダーと比較してデータ品質 (欠損、遅延、精度) に制約がある。重要な投資判断においては、Bloomberg 連携による高品質データでのクロスバリデーションを実施する。

### 10.3 規制・コンプライアンスリスク

**AI 投資判断の説明責任:** 金融庁が求める投資判断の説明責任に対し、AI の判断プロセスの透明性確保が課題となる。エージェント間の議論ログを監査トレースとして永続化し、自然言語で判断根拠を遡及説明可能にする HITL 設計で対応する。ただし、LLM の内部推論過程自体は依然としてブラックボックスであり、「なぜその文章を生成したか」の完全な説明は原理的に困難である。



**金融規制への準拠:** AI を活用した投資判断が拡大する中、今後の規制強化（EU AI Act の金融分野への適用拡大等）に備え、Human-in-the-Loop による最終判断権限の人間への留保と、判断プロセスの記録・保存を設計段階から組み込む。

**データプライバシー:** アナリストレポートやファンドマネージャーの判断データは機密情報である。Claude API への送信データの範囲を最小限に制御し、ローカルで処理可能な計算はローカルで完結させる設計を徹底する。

# 付録

## A. 参照論文一覧

#	タイトル	著者/機関	年	主要貢献
1	AlphaAgents: Large Language Model based Multi-Agents for Equity Portfolio Constructions	BlackRock	2025	Fundamental/Sentiment/Valuation の3専門エージェントが Round Robin 議論を通じて投資判断を形成。Financial RAG Tool による10-K解析
2	MASFIN: A Multi-Agent System for Decomposed Financial Reasoning and Forecasting	Rochester Institute of Technology	2025	Postmortem→Screening→Analysis→Timing→Portfolio の5段階パイプライン。8週間のライブ検証で7.33%リターン
3	TradingAgents: Multi-Agents LLM Financial Trading Framework	UCLA / MIT	2025	Bull/Bear の弁証法的アーキテクチャ。Risk Management Team によるバイアス検出。AutoGen 基盤
4	HedgeAgents: A Balanced-aware Multi-agent Financial Trading System	South China University of Technology	2025	バランス重視のリスク管理。Conservative/Aggressive/Neutral の3タイプエージェント。数理最適化統合
5	FinRobot: An Open-Source AI Agent Platform for Financial Applications using Large Language Models	AI4Finance Foundation	2024	オープンソースプラットフォーム。マルチソース統合。強化学習によるポートフォリオ最適化
6	StockBench: Can LLM Agents Trade Stocks Profitably In Real-world Markets?	Tsinghua University	2025	リアル市場でのLLMエージェント検証。データ汚染チェック手法の提案。バックテストと実運用のギャップを定量評価
7	Automate Strategy Finding with LLM in Quant Investment	HKUST	2024	Seed Alpha の LLM 自動生成。9カテゴリー100因子。SSE 50 において53.17%累積リターン（2023年1月-2024年1月）
8	Can LLM-based Financial Investing Strategies Outperform the Market in Long Run?	University of Edinburgh	2025	長期パフォーマンス評価フレームワーク FINSABER。市場レジーム変化への適応性を検証

#	タイトル	著者/機関	年	主要貢献
9	Assessing Look-Ahead Bias in Stock Return Predictions Generated By GPT Sentiment Analysis	Glasserman, Li, Mamaysky (Columbia)	2024	GPT によるセンチメント分析における先読みバイアスの定量的評価。時間的整合性の重要性を実証

B. Y のスコアリング体系（全表）

B.1 確信度5段階

ランク	確信度	日本語表現	判断基準	代表例
かなり納得	90%	構造的優位性 + 明確な CAGR 接続 + 定量的裏付け	因果チェーンが明確、数値比較あり、構造的（一時的でない）	COST 顧客粘着力（更新率90%超 + ARPU向上メカニズム）
おおむね納得	70%	妥当な仮説 + 一定の裏付け	方向性が妥当で一定の証拠あり、ただし定量性に改善余地	CHD #1 買収能力（買収基準明確化 + 人員投入）
まあ納得	50%	方向性は認めるが裏付け不十分	論理に飛躍はないが、定量的証拠が乏しい	一般的な競争優位性の主張で部分的根拠あり
あまり納得しない	30%	飛躍的解釈・因果関係の逆転	結果を原因と取り違え、一般論、戦略と優位性の混同	MNST #5（低所得→コンビニ→粘着力の飛躍）
却下	10%	事実誤認・競争優位性として不成立	明白な事実誤認、論理的に成立しない	MNST #6（OPM を GPM と誤認）

B.2 却下パターン7類型

#	パターン	説明	具体例
1	結果を原因と取り違え	観測される結果を競争優位性の原因として提示	「高シェアは結果であり背後要因が乏しい」（MNST #1）
2	どの企業にも言える一般論	業界内で広く共通する特徴を固有の優位性として主張	「多くの企業が言っていること」（COST #5）
3	競合との差別化が不明	競合他社との具体的な差異が示されていない	「純粋競合に対して優位な設計かは不明」（COST #3）
4	事実誤認	財務指標や事業データの誤引用	「OPM を GPM と誤認」（MNST #6）
5	飛躍的解釈	論理チェーンに根拠のない飛躍が含まれる	「低所得→コンビニ購買率高い→粘着力高い」（MNST #5）
6	定性的で定量的裏付けなし	主張に具体的な数値データが伴わない	「定性的差異がわかりにくく定量評価もしがたい」（COST #2）
7	戦略を優位性と混同	実行可能な戦略を参入障壁と混同	「ドミナント出店戦略＝戦略は優位性ではない」（ORLY）

B.3 高評価パターン6類型

#	パターン	説明	具体例
---	------	----	-----

#	パターン	説明	具体例
1	定量的裏付け	具体的な数値比較を伴う主張	退職率9% vs 業界20%+ (COST #1)
2	直接的 CAGR 接続メカニズム	優位性から CAGR への因果経路が明示	上級会員ミックス改善→ARPU向上→既存店売上成長 (COST)
3	能力 > 結果	プロセス・能力に着目し、結果指標ではない	ポート管理能力が優位性 (CHD #3)
4	明確な因果チェーン	複数ステップの論理が各段階で根拠を持つ	買収基準明確化→人員投入・精査→選定能力 (CHD #1)
5	他社事例との類推	過去の判断事例や他社との比較による裏付け	ブランド力評価で過去判断事例を参照 (MNST #1)
6	構造的 vs 補完的の区別	CAGR 寄与を構造的要素と補完的要素に明確に分類	COST ガソリン戦略の位置づけ

## C. パフォーマンス指標の数式一覧

指標	数式	説明
年率リターン (Annualized Return)	$R_{\text{ann}} = \left( (1 + R_{\text{cum}})^{252/n} \right) - 1$	累積リターン $R_{\text{cum}}$ を取引日数 $n$ で年率換算
年率ボラティリティ (Annualized Volatility)	$\sigma_{\text{ann}} = \sigma_{\text{daily}} \times \sqrt{252}$	日次標準偏差を年率換算
シャープレシオ (Sharpe Ratio)	$S = \frac{R_p - R_f}{\sigma_p}$	ポートフォリオリターン $R_p$ からリスクフリーレート $R_f$ を差し引き、ボラティリティ $\sigma_p$ で除算
ソルティノレシオ (Sortino Ratio)	$S_{\text{sort}} = \frac{R_p - R_f}{\sigma_{\text{downside}}}$	下方リスク $\sigma_{\text{downside}}$ のみを考慮したリスク調整後リターン
カルマーレシオ (Calmar Ratio)	$C = \frac{R_{\text{ann}}}{\text{MDD}}$	年率リターンを最大ドロウダウン MDD で除算
情報係数 (Information Coefficient)	$\text{IC} = \text{Corr}(\hat{r}, r)$	予測リターン $\hat{r}$ と実績リターン $r$ のピアソン相関係数
最大ドロウダウン (Maximum Drawdown)	$\text{MDD} = \max_t \left( \frac{\text{Peak}_i - \text{Trough}_t}{\text{Peak}_i} \right)$	ピークからトラフまでの最大下落率
VaR (Value at Risk)	$\text{VaR}_\alpha = -F^{-1}(\alpha)$	信頼水準 $\alpha$ における最大損失額。 $F^{-1}$ は損益分布の逆関数
CVaR (Conditional VaR)	$\text{CVaR}_\alpha = -E[X \mid X \leq -\text{VaR}_\alpha]$	VaR を超過した場合の期待損失額
アルファシグナル統合式	$\alpha_{ij}^{(t)} = w_{\text{cs}} \cdot f_{\text{cs}}(X_i^{(t)}) + w_{\text{ts}} \cdot f_{\text{ts}}(X_i^{(t)}, \dots, X_i^{(t-n)})$	クロスセクション演算子 $f_{\text{cs}}$ と時系列演算子 $f_{\text{ts}}$ の加重合成

指標	数式	説明
Black-Litterman 期待リターン	$E[R] = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1} [(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q]$	均衡リターン $\Pi$ 、ビュー行列 $P$ 、ビューリターン $Q$ 、ビュー不確実性 $\Omega$

D. 用語集

D.1 金融用語

用語	定義
アルファ（Alpha）	市場ベンチマークを上回る超過収益。ポートフォリオマネージャーの付加価値を測定する指標
CAGR（Compound Annual Growth Rate）	年平均成長率。投資期間全体にわたる平均的な年間成長率
ファクター投資（Factor Investing）	バリュー、モメンタム、クオリティ等の体系的なリスクプレミアムに基づく投資手法
バリュエーション（Valuation）	企業の本源的価値（Intrinsic Value）の評価。DCF、マルチプル等の手法を包含
シャープレシオ（Sharpe Ratio）	リスク1単位あたりの超過リターン。リスク調整後パフォーマンスの標準的指標
ソルティノレシオ（Sortino Ratio）	下方リスクのみを考慮したリスク調整後リターン指標
最大ドローダウン（Maximum Drawdown）	ピークからトラフまでの最大累積損失率。リスク耐性の指標
Black-Litterman モデル	市場均衡リターンと投資家のビュー（主観的予測）を統合するポートフォリオ最適化手法
平均分散最適化（Mean-Variance Optimization）	Markowitz の現代ポートフォリオ理論に基づく、期待リターンと分散を最適化する手法
10-K / 10-Q	SEC に提出する年次報告書（10-K）および四半期報告書（10-Q）
8-K	SEC に提出する臨時報告書。重要事象（M&A、経営陣交代等）の開示
生存者バイアス（Survivorship Bias）	上場廃止・破綻した銘柄を分析対象から除外することで生じる上方バイアス
先読みバイアス（Look-ahead Bias）	分析時点では利用不可能な将来の情報を無意識に使用するバイアス
センチメント分析（Sentiment Analysis）	ニュース、SNS、アナリストレポート等のテキストから市場心理を定量化する手法
インサイダー取引（Insider Trading）	企業内部者による自社株の売買。経営陣の確信度の指標として分析される

用語	定義
競争優位性 (Competitive Advantage)	競合他社に対する持続的な優位性。本プロジェクトでは Y の判断軸に基づき厳密に評価
7 Powers	Hamilton Helmer のフレームワーク。Scale Economies, Network Effects, Counter-Positioning, Switching Costs, Branding, Cornered Resource, Process Power の7つの持続的競争優位の源泉
モメンタム (Momentum)	過去のリターンが将来のリターンを予測するアノマリー。上昇トレンドの継続性
ボラティリティ (Volatility)	価格変動の標準偏差。リスクの代理変数として使用

D.2 AI/ML 用語

用語	定義
LLM (Large Language Model)	大規模言語モデル。Transformer アーキテクチャに基づく大規模パラメータの言語生成モデル
MAS (Multi-Agent System)	マルチエージェントシステム。複数の自律エージェントが協調・対立して目標を達成する計算システム
ハルシネーション (Hallucination)	LLM が事実に反する情報を生成する現象。金融分析では数値の誤引用として顕在化
RAG (Retrieval-Augmented Generation)	検索拡張生成。外部知識ベースから関連文書を検索し、LLM の生成精度を向上させる手法
Few-shot Learning	少数の例示を与えることで LLM の出力パターンを制御する手法
プロンプトエンジニアリング (Prompt Engineering)	LLM の入力テキスト (プロンプト) を最適化し、望ましい出力を得るための設計技法
システムプロンプト (System Prompt)	LLM の振る舞いを定義する初期指示。エージェントのペルソナと行動規範を規定
Round Robin	エージェント間の議論で、各エージェントが順番に発言・レビューを行う方式
Reflection (自己反省)	エージェントが自身の過去の出力を評価・修正するメタ認知プロセス
MLP (Multi-Layer Perceptron)	多層パーセプトロン。入力層・隠れ層・出力層からなるフィードフォワードニューラルネットワーク
確認バイアス (Confirmation Bias)	既存の信念を支持する情報を選好し、矛盾する情報を軽視する認知バイアス
Seed Alpha	LLM が自動生成するアルファファクターの候補。人間のクオンツが設計する従来のファクターに対応
ベクトル DB (Vector Database)	高次元ベクトル (埋め込み表現) の類似検索に特化したデータベース。RAG の基盤技術
Agent Teams	Claude Code における複数エージェントの協調実行フレームワーク

D.3 システム用語

用語	定義
Claude Code	Anthropic が提供する自律型開発 CLI。エージェント定義・スキル・コマンドによるオーケストレーション基盤
HITL (Human-in-the-Loop)	AI システムの判断プロセスに人間の検証・承認を組み込む設計パターン
Audit Trace (監査トレース)	判断プロセスの全記録。規制要件への対応と事後分析のための証跡
Temporal Alignment (時間的整合性)	分析時点のデータのみを使用し、将来情報の混入を防止する設計原則
Graceful Degradation (グレースフルデグラデーション)	外部サービス障害時にシステム全体の停止を防ぎ、機能を段階的に縮退させる設計
TDD (Test-Driven Development)	テスト駆動開発。Red→Green→Refactor のサイクルで品質を担保する開発手法
uv	Rust 製の高速 Python パッケージマネージャー。pip/venv の代替
yfinance	Yahoo Finance の非公式 Python API。株価・財務データの取得に使用
FRED (Federal Reserve Economic Data)	米連邦準備制度が提供する経済データベース。マクロ経済指標の取得に使用
edgartools	SEC EDGAR のデータを Python で取得・解析するライブラリ
DuckDB	分析特化型の組み込みデータベース。大規模データの高速集計に使用
Ruff	Rust 製の高速 Python リンター・フォーマッター
pyright	Microsoft 製の Python 静的型チェッカー
Hypothesis	Python のプロパティベーステストフレームワーク。ランダムデータによる不変条件検証