



次世代金融研シリーズ

金融化 is eating the world. —AI で加速する 「金融化」と信頼アーキテクチャー

山上 聡（NTT データ経営研究所 フェロー）*

オーケストレーションの先にある問い

前稿の「[断片化する金融システムを束ねる『オーケストレーター』としての銀行](#)」（山上, 2026）では、断片化する通貨、台帳、金融機能を束ねるオーケストレーターとして、銀行が果たし得る役割を論じた。しかし、AI の包括的な発展により、経済活動の相当割合が「金融が処理可能な情報活動」に転じていく段階が、そう遠くない将来に訪れよう。そこでは、オーケストレーターが束ねる対象は、既存の金融商品や決済手段にとどまらない。情報産業としての広義金融サービスが経済活動や意思決定を飲み込んでいく事象が生じるであろう。そうした未来像を想定すれば、次に問われるのは、未確定の未来社会をどのように広義金融サービスの範疇に取り込んでいくか（後述する金融化をどう実現するか）である。そこでは、こうしたビジネスモデルの変革が社会に安心して受け入れられるための信頼アーキテクチャー（Trust Architecture）をどう築いていくかが大きな課題となる。

デジタル化と AI の進展を金融機関の業務効率化として捉えるだけでは、この変化の大きさを見誤る。紙や会話をデータに変える Digitization、データによって既存業務を連結・自動化する Digitalization を経て、金融の境界、役割、主体間関係を組み替える段階に入ると、金融は経済活動の外側から商品を提供する存在ではなく、経済活動の内部で意思決定を支援・実行するサービスの提供主体となる。本稿では、この変化を「金融化」という視点から捉え、その具体的な事例として、利用者側の AI であるロボ CFO と、高度なサイバー能力を持つ AI 「Claude Mythos」が、将来の金融に必要な信頼の二つの側面をどのように浮かび上がらせるかを考えたい。

デジタル化が拡張する「金融化」の範囲

金融化という言葉は、実体経済に対する金融部門や株主価値の過度な影響を批判する意味で用いられることがある。本稿では、これとは異なり、経済活動がデジタルに観測され、将来キャッシュフロー（CF）、リスク、権利として評価・価格付け・資金配分の対象になっていくことを指す。

従来の金融システムがデジタル化してきた対象は、主として金融機関内部の確定したビジネスであった。振込、預金、約定、決済、返済などは高度にシステム化されている。一方、それらの金融取引の背景にある経済活動、すなわち企業が何を実現しようとしているかを捉え、これを金融サービスの高度化に繋げていく仕組みとインフラ基盤は、未だ十分にはデジタル

化されていない。受注の見込み、契約交渉、履行の進捗、要員の確保、顧客との関係、設備投資の目的といった情報は、メール、文書、表計算ソフト、会話、そして経営者の頭の中に残されている。

金融機関に資金需要が現れるのは、多くの場合、こうした経営判断の結果として資金不足が顕在化した後である。しかし資金需要は、融資の申込みによって突然生まれるわけではない。人を採用する、案件を受注する、新事業へ投資する、既存事業を継続するといった利用者側の判断から資金需要は始まっている。生成 AI が文書や会話から意図と前提を読み取り、異なるデータを結び、複数の将来シナリオを比較できるようになると、この「金融の手前」をコンピュータが扱えるようになる。

その結果、利用者側の金融と金融機関側の金融は次第に一体化する。企業が最初から融資、保険、増資、投資、M&A、事業承継などの商品サービスを探すのではなく、まず事業上の目的と制約を検討し、それでも残る資金・リスク上の制約に対して金融機能を組み合わせる。金融商品が検討の出発点ではなく、経済活動を実現するための構成要素になるのである。これが、AI によって拡大する金融化の本質である。

ロボ CFO——未来を予言せず、実現過程を観測する

この考え方を利用者側から具体化する試みが、筆者らが現在開発・検証している「ロボ CFO」である。ロボ CFO は、採用、受注、投資、事業継続などについて、経営者の目的を問い返ししながら構造化し、案件、契約、要員、コスト等の情報から将来 CF へ翻訳する。複数の選択肢を比較し、事業上の対応を講じてもなお残る資金制約を特定して、必要な金融へ引き渡す。

ここで重要なのは、将来 CF の予測精度だけではない。将来 CF は、予測した時点では確率的な期待にすぎない。しかし、商談が契約となり、契約が履行され、納品・検収を経て請求可能な債権となり、最後に入金されることで確定した実績へ変わる。厳密には、確率が一挙に確定へ変わるのではなく、経済活動の進展と証拠の蓄積に伴って確定度を高めていく。ロボ CFO が観測すべきなのは、将来を言い当てたか否かだけではなく、前回の予測から何が確定し、何が未確定のまま残り、どの前提が変わったかという状態の遷移である。

言い方を変えれば、これらは経営者や社員の頭の中にある企業活動のオントロジーを整理・明確化する行為である。それらは意味関係の構造化であり、これを通じて数値となり、企業財務や金融サービスと一体化していく。AI の活用においては、こうしたオントロジーを言語化する過程が必要となる。これらが暗黙知のままでは AI は機能しない。企業の情報システムへのアクセスだけでは足りないのである。

筆者がかつて米国監査法人で勤務していた時代、「信頼は主張ではなくプロセスによって証明するものだ」と教えられた。誰が、どの情報に基づき、どの手続を行い、誰が確認・承認し、どの証拠が残り、例外がどのように処理されたかを追跡可能にすることで、信頼は形成される。この原則を将来 CF に適用すれば、未来を事前に「証明」する必要はない。未来について

の主張（一定の確度を持った見通し）が、契約、履行、検収、請求、入金というプロセスによって検証され、確定事実へ近づいていく道筋を示せばよい。

ただし、企業が進捗率や入金予定を自ら入力しただけでは、自己申告をデジタル化したにすぎない。「証拠の強さ」も同時に評価しなければならない。担当者の申告、社内の案件・稼働記録、取引相手による発注・検収、電子契約、物流、銀行入金などでは、証拠としての独立性と確度が異なる。これらを照合し、当初予測と実績の差異を継続的に蓄積すれば、その企業が「約束を実現する能力」も見えてくる。ここにビッグデータとAIのパワーを活用するのである。従来の決算書が過去の結果を一定時点で示すものだとすれば（BPOやSaaSがこれを支援）、ロボCFOは将来が実現する過程を動的に示すものとなる。筆者は、このより広い構想を、将来CFの実現過程と金融機能をつなぐ「EXPECTUS」（Expectation into Reality）と呼ぶことを検討している。

Claude Mythos——証明プロセス自体を信頼できるか

ロボCFOが「何を観測すれば経済的な信頼になるか」を示すとすれば、Claude Mythosは「その観測と証拠を支えるデジタル基盤を本当に信頼できるか」という問いを突きつける。

Anthropicが2026年4月に公表したClaude Mythos Previewは、高度なコーディングと推論能力を背景に、主要なOSやブラウザに存在する未知の脆弱性を発見し、条件によってはその悪用方法まで構築できる能力を示した。同社は一般公開を行わず、重要ソフトウェアを防御するProject Glasswingの参加組織に限定して提供した（Anthropic, 2026a）。初期報告では、Claude Mythosが大量の高深刻度の脆弱性候補を発見する一方、人間による検証、開示、修正、パッチ適用が新たなボトルネックになったことが示されている（Anthropic, 2026b）。これは単に攻撃AIが強力になったという話ではない。脆弱性の発見と攻撃が機械速度になる一方、確認と修復は人間と組織の速度に制約されるという、時間の非対称性が生じたのである。

金融にとってさらに重要なのは、共通ソフトウェア、クラウド、決済ネットワーク、外部委託先を通じて、同じ脆弱性が多数の機関に同時に波及し得る点である。IMFは、AIがもたらす金融安定上の問題は、必ずしも未知の攻撃類型にあるのではなく、既存の攻撃を高速・大規模・同時多発化させる「規模の効果」にあると指摘している（Adrian et al., 2026）。日本でも金融庁と日本銀行が2026年5月、フロンティアAIによる脅威変化を踏まえ、金融機関等に短期的対応を要請した（金融庁・日本銀行, 2026）。

Claude Mythosは脅威であると同時に、防御にも利用できる。その意味で「悪いAI」ではない。むしろ象徴しているのは、AIが攻撃側と防御側の能力を同時に引き上げ、従来の安全性の前提を崩したことである。ロボCFOが契約、履行、検収、請求、入金を精緻に追跡しても、元システムが侵害され、本人や取引相手が偽装され、証拠が改ざんされ、AIの判断が誘導されていれば、その観測は金融上の信頼を生まない。また、正しい判断であっても、AIエージェ

ントに過大な実行権限が付与されていれば、誤作動や乗っ取りが即座に資金移動へ波及する。

二つの信頼を結ぶ Trust Architecture

ここから、将来の金融に必要な信頼は二層に分けて考えられる。第一は経済的信頼である。約束された将来 CF が、経済活動の進展によってどこまで実現に近づいているかという判断であり、ロボ CFO が主として扱う。第二は技術的・制度的信頼である。その観測主体、データ、証拠情報、AI 判断、取引実行が真正であり、異常時に被害を限定し、かつ円滑に回復できるかという問いであり、Claude Mythos がその重要性を浮かび上がらせた。

したがって、金融上の信頼は、概念的には「将来 CF の実現プロセスの可視性」「観測・証拠の真正性」「判断・実行の統制と回復可能性」の掛け算として捉えられる。どれか一つが欠ければ、全体の信頼は成立しない。経済活動が順調でもデータが偽造されていれば信頼できず、データが真正でも将来 CF の実現過程が見えなければ資金配分を判断できない。

ここでいう Trust Architecture は、従来型のサイバーセキュリティを追加することではない。データの発生源と来歴、本人・企業・AI エージェントのアイデンティティ識別、最小権限、判断と実行の分離、高額・不可逆取引への複数承認、継続的な検証、異常時のサーキットブレーカー、縮退運転、復旧、損失負担と責任分担までを含む。金融化によって観測可能な範囲が広がり、相互運用によって接続先が増え、エージェント化によって判断と実行が自動化されるほど、Trust Architecture は後工程ではなく、金融化そのものの成立条件になる。

銀行は「唯一の指揮者」から「信頼のアンカー」へ

前稿では、銀行が断片化した金融を束ねるオーケストレーターへ進化する必要性を論じた。続編となる本稿の検討を踏まえると、その役割をさらに精緻に捉える必要がある。将来のオーケストレーションは、一つの主体がすべてを支配する形にはならない可能性が高い。企業側の AI や業務 SaaS が利用者の目的を理解し、金融プラットフォームが融資、証券、保険、保証、VC 等を組み合わせ、銀行や決済基盤が本人確認、規制対応、資金移転、記録、損失負担、危機対応を担うという分業が考えられる。

したがって、銀行が唯一のオーケストレーターになるとは限らない。非金融業であるエコシステムプレーヤーや情報インフラ企業なども、その候補たり得る。しかし、銀行には顧客確認、預金、決済、与信、規制対応、補償、危機管理を長年担ってきた制度的・実務的な蓄積がある。これを閉じた商品・サービス提供のために使うのではなく、外部の AI、SaaS、証券、保険、ノンバンク、VC 等が参加する金融サービス形成の中で、真正性と実行の最終的な拠り所として提供できれば、銀行は強い「信頼のアンカー」になり得る。

これは前稿のオーケストレーター論を後退させるものではない。オーケストレーションの対象を、金融商品や価値媒体だけでなく、将来 CF、証拠、権限、責任、停止・回復まで広げるものである。銀行が顧客接点を独占することではなく、利用者側で形成された目的と金融ニー

ズを安全に受け止め、複数の主体が構成するサービスに信頼を与えることこそ、AI時代の銀行の中核的役割となる。これが本稿が主張する「金融化」の意味するところであり、これを成し得た金融業は「金融化 is eating the world.」の具現者となろう。

おわりに——未来の実現過程を金融にする

従来の金融は、過去の決算、企業属性、返済実績、担保・保証によって、直接には見えない未来を補ってきた。これらの重要性は今後も失われない。しかし、経済活動がデジタルに観測可能になれば、未来についての約束が契約、履行、検収、請求、入金を経て確定事実へ変わるプロセスを、金融判断に加えることができる。

AIが金融にもたらす最大の変化は、既存業務を速く安くすることだけではない。これまで金融の外側にあった経済活動を理解し、将来CFとして表現し、その実現過程を観測しながら、必要な金融機能を組み替えることである。ロボCFOは未来の実現プロセスを信頼へ変える可能性を示し、Claude Mythosが象徴するAIのパワーはそのプロセス自体を信頼可能にする設計の必要性を示している。

金融とは、未来に対して現在の価値を渡す仕組みである。その未来を信頼する方法が、過去の属性と失敗時の保全だけでなく、真正に観測された実現プロセスへ広がるとき、金融化は単なる金融ビジネスの対象範囲の拡大ではなく、企業と金融の関係そのものの再設計となる。オーケストレーションの時代は、同時にTrust Architectureの時代なのである。

参考文献

金融庁・日本銀行（2026）「フロンティア AI による脅威変化を踏まえた金融機関等の短期的な対応」2026年5月22日。

山上 聡（2026）「[断片化する金融システムを束ねる『オーケストレーター』としての銀行](#)」SBI金融経済研究所、レポート、2026年5月25日。

Adrian, T. et al. (2026) “Artificial Intelligence and Cybersecurity in the Financial Sector,” IMF Note 2026/005.

Anthropic (2026a) “Assessing Claude Mythos Preview’s cybersecurity capabilities,” April 7, 2026.

Anthropic (2026b) “Project Glasswing: An initial update,” May 22, 2026.

* 筆者の山上聡氏（NTTデータ経営研究所 フェロー）は、SBI金融経済研究所が主催する「[次世代金融インフラの構築を考える研究会](#)」のメンバーである。レポートの次世代金融研シリーズは、同研究会が扱うテーマに関連したレポートをシリーズ化するものであり、本レポートが初号となる。