FUJITSU

AIファースト時代の金融業変革

「AIを使う」から「AIを経営する」へ



Contents

AIファースト時代の金融業変革：「AIを使う」から「AIを経営する」へ

はじめに
AIファースト時代、金融業変革の新たなステージ …… 3

1. AIエージェントを「導入する」から「組織に組み込む」へ
―組織・ワークフローの再設計 …… 3

2. AIに「タスクをさせる」から「役割を担わせる」へ
―金融業におけるアプローチ：
AIアシスタントからデジタル労働力へ …… 6

3. AIを「使う」から「経営する」へ
―金融業における実践：AI価値の最大化とトークン（Token）経営 …… 8

4. AI時代のガバナンス：金融業の新たな競争力
―リスク対応から「AIレジリエンス」へ …… 12

おわりに
AI経営に向けて
―金融業リーダーへの3つの提言 …… 14



はじめに

AIファースト時代、金融業変革の新たなステージ

生成AI、AIエージェントの進化を背景に、金融機関のAI活用は、個々の組織メンバーの生産性向上や個別業務の効率化から、組織・業務・顧客接点、さらにはビジネスモデルそのものを変革する段階へと移行しつつあります。

特に先進的な金融機関では、AIを前提としてワークフローや組織構造を再設計するとともに、人間とAIが協働するための職種・役割を再定義し始めています。さらに、AIを新たな経営資源として捉え、投資対効果を定量化しながら、価値創出を最大化するためにAIをいかに「経営するか」という取り組みも始まっています。金融業の変革は、いま新たなステージに入りつつあります。

一方、AIインテリジェンスの急速な進化に伴い、自律的なタスク実行能力が大きく高まる一方で、一部のフロンティアAIでは、ガードレールの回避や外部システムへの不正アクセスなど、新たなリスクも顕在化しています。また、高度なAIモデルが悪用される可能性も懸念されています。こうしたリスクを適切に管理・低減し、能動的なAIリスク管理とガバナンスを、AI活用を支えるだけでなく**企業競争力の一部として確立すること**が求められています。

本稿では、こうした変化を「AIファースト」という視点から捉え、**技術の最適化から組織・役割の再設計、AI経営、そして新たな競争力の形成まで**、金融業に求められる経営変革の方向性を考察します。

1. AIエージェントを「導入する」から「組織に組み込む」へ

―組織・ワークフローの再設計

KPMGの「AI Quarterly Pulse Survey Banking 2026 Q2」によると、金融業ではAIエージェントの活用が、個別タスクから複数の組織機能やワークフローへと広がりつつあります。[*1] 特に、複数の組織機能へAIエージェントを展開・スケールする段階が24%、複数のAIエージェントをワークフロー横断で協調・統合する段階が15%に達しています。AIエージェントは、個別のユースケースから、組織的な業務遂行へと活用範囲を広げ始めています。

*1 KPMG (June 2026) "AI Quarterly Pulse Survey — Banking Q2 2026"

AIを組織へ統合する進化プロセス

AIエージェントは、単独でタスクを遂行するだけでなく、複数のエージェントが役割を分担したり、人間や外部システムと連携したりすることができます。技術的には、Single Agent、Multi-Agent、Agent Ecosystemへと発展しています。

ただし、重要なのは、こうした技術的な構成と、組織におけるAIの役割は同一ではないことです。**Single Agentであっても、与えられる役割・責任・権限によって、その組織上の位置づけは変わります。**AI活用の進化を捉えるには、技術だけでなく、AIが組織の中で「何を担うのか」を見る必要があります。

表1は、金融機関におけるAI活用の進化を、**AIが担う役割と組織への統合レベル**という観点から整理したものです。

表1　AI活用の進化と組織への統合

項目	AIの役割	人間との関係	典型的な応用例
① **AIアシスタント**	人間の仕事を支援	人間主導	Q&A、要約、情報支援
② **AIエージェント**	人間から与えられた目的に基づきタスクを実行	目標志向	調査、分析、多段階のタスク実行
③ **デジタルワーカー**	組織の中で特定の役割を持ち、業務を遂行	役割ベース	ワークフローやシステムを横断した、役割ごとの業務
④ **デジタル労働力**	複数のデジタルワーカーが連携し、組織的な業務を遂行	組織レベル	ワークフロー、機能、あるいは組織の境界を越えて複数のAIエージェントが連携して作業遂行

出所：著者作成

① AIアシスタント（人間支援型）

AIアシスタントは、人間の仕事を支援します。LLM、RAG、Web検索などの進化により、情報検索・分析から複数ステップのタスクまで対応範囲が広がっています。そのため、AssistantとAgentの技術的な境界は必ずしも明確ではありません。重要なのは技術名称ではなく、**人間を支援する役割なのか、AI自身が目的に向けて業務を遂行するのか**という違いです。

② AIエージェント（目標志向型）

AIエージェントは、人間が設定した目的や条件に基づき、必要な情報やツールを選択しながら複数のステップを自律的に実行します。AIは「回答する存在」から、**目的に向かってタスクを実行する存在**へと役割を広げます。

③ デジタルワーカー（役割ベース型）

デジタルワーカーは、組織の中で特定の役割を担い、継続的に業務を遂行するAIです。責任範囲、権限、KPI、意思決定範囲などを設定し、複数のシステムやワークフローを横断して業務を遂行できます。ここでは、AIを単なるツールではなく、**組織の中で役割を担う労働力として位置づける**ことがポイントになります。

④ デジタル労働力（組織連携型）

デジタルワークフォースは、複数のデジタルワーカーが役割を分担・連携し、組織レベルの業務を遂行する段階です。Sales、KYC・KYA、Compliance、Credit、Operationsなど、複数の機能をまたぐエンドツーエンドの業務への展開が想定されます。さらに将来は、**外部組織のAIエージェントとも、KYA（Know Your Agent）などの信頼基盤のもとで連携するAgent Ecosystem**へ発展する可能性があります。

このように、AI活用の進化は単なる技術性能の向上ではありません。

AIアシスタント → AIエージェント → デジタルワーカー → デジタルワークフォースという流れは、AIが「人間を支援する」→「タスクを実行する」→「組織の中で役割を担う」→「複数のAIが連携して業務を遂行する」という、役割と責任の拡大を意味します。2025年にPoCや個別導入が広がったのに対し、2026年は、AIエージェントを個別ユースケースにとどめず、**業務・組織に組み込み、スケールさせる段階への移行期**を迎えています。

AIエージェント時代へ適応する組織変革：機能別組織から「ワークチャート型組織」へ

AIエージェントが組織に組み込まれるにつれ、問われるのはAIをどう導入するかだけではなく、**AIを前提に仕事そのものをどう再設計するか**です。

これまで企業は、AIの効果をイノベーション、業務改革、顧客体験、生産性向上といった機能別の成果として捉えてきました。今後は、**エンドツーエンドの価値創造**を起点に、従来の機能別組織から「**ワークチャート型組織**」への転換が進むと考えられます。従来の組織図が「誰が誰を管理するか」を示すのに対し、ワークチャートは「何を、誰が、どのように実行するか」を起点とします。業務・タスク・アウトカムの構造に沿って、人とAIを動的に配置・再接続する組織モデルです。人の専門性を起点に仕事を割り当てる機能別組織に対し、ワークチャート型組織は、**仕事そのものを起点に人とAIの役割を再設計する**点に特徴があります。

その結果、組織はより流動的になり、部門の境界を越えて人とAIの能力を組み合わせることで、**組織サイロを解消し、エンドツーエンドの価値創造を加速する**ことが期待されます。



2. AIに「タスクをさせる」から「役割を担わせる」へ
―金融業におけるアプローチ：AIアシスタントからデジタル労働力へ

第1章で述べたように、金融業ではAIエージェントの活用領域が急速に拡大しています。顧客接点・営業支援、営業・市場分析、クレジット審査、バックオフィス、IT・ソフトウェア開発、監査・コンプライアンスなど、**活用領域の拡大に加え、AIが担う仕事も深化**し始めています。

従来の個別タスクの支援・実行から、複数のステップやシステムをまたぐワークフローを遂行し、組織の中で一定の役割を担うAIへと進化しています。JPMorgan Chaseが指摘するように、AIエージェントは単一タスクを処理するツールから、複数ステップのワークフローを担う「**デジタルワーカー（Digital Worker）**」へと発展しつつあります。[*2]

今後は、AIを単に「使う」のではなく、**組織の役割を担うデジタルワーカーとして組み込み、さらに複数のデジタルワーカーを連携させるデジタルワークフォースへ発展させること**が、次の実装段階となります。

先行事例に見るデジタルワーカーとデジタルワークフォース

表2は、AIが「役割を担う」段階から、**組織内、さらには企業間で協働する段階**へ進む5つの先行事例を示したものです。

表2　金融業界におけるAIの「労働力化」からAgentic Ecosystemへの先行5事例

企業名	労働力化プロジェクト	主な応用	特長	示唆
Goldman Sachs（米国）	AIデジタルコワーカー	会計・コンプライアンスなどのバックオフィス	AIを「デジタルな同僚」として活用。複雑・ルール型業務を遂行	**タスク特化型**
三菱UFJ銀行（日本）	「AI行員」構想	スピーチ作成など20超の社内業務	AIをデジタルな「銀行行員」として位置づけ。人材管理の考え方をAIに応用	**人材管理統合型**
BNY Mellon（米国）	デジタル従業員	支払指図の検証、コード開発・修正など	AIにID・権限・上司を付与し、組織管理に組み込む	**組織統合型**
Morgan Stanley（米国）	AIエージェントと自社PFの連携	Shareworks/Equity Edgeへの顧客AIエージェント接続	人間向けUIを介さず、AIエージェントが金融PFと直接連携	**Agentic Interface型**
Mastercard（米国）	Agent Pay/Agent Suite	AIによる商品探索・購買・決済、企業業務	Agentの登録・認証、権限管理、上限設定、トークン化などを統合	**Agentic Ecosystem型**

出所：各社のプレスリリース、報道などにより、著者作成

*2　CNBC（June 9, 2026）“JPMorgan Chase plans to deploy more powerful AI agents this year”

表2のまとめ・示唆

表2のGoldman Sachs、三菱UFJ銀行、BNY Mellonの3社は**社内でAIをデジタルワーカーとして実装する事例**であるのに対し、Morgan Stanleyは**顧客AIエージェントと自社プラットフォームを接続する事例**、Mastercardは**顧客・金融機関間の自律的な取引を支える基盤を提供する事例**です。

Goldman Sachsは、会計・コンプライアンス・顧客手続きなどのバックオフィス業務でAIを「Digital Coworker」と位置づけ、複雑でルール集約型の業務への適用を進めています。[*3] **三菱UFJ銀行**は、AIを「AI行員」として採用・育成・配置する人事モデルのもと、スピーチ作成や社内問い合わせ対応など複数業務へ展開しています。[*4] **BNY Mellon**は、AIを「仮想従業員」として組織に組み込み、権限設定や実績・品質管理など、人間の従業員に近い管理モデルを導入しています。[*5] このように、2025年後半以降、大手銀行では、ミドル・バックオフィスを中心に、AIエージェントを**デジタル・コワーカー/デジタルワーカーとして業務プロセスに組み込む動き**が加速しています。

さらに、**Morgan Stanley**は、顧客企業のAIエージェントが株式プラン管理プラットフォームに直接アクセスできるAgent-to-Platform連携を実装しています。[*6] これは、従来の「顧客 → Web/アプリ → サービス」という顧客接点が、**「顧客企業Agent → 金融プラットフォーム → サービス」へ転換する可能性**を示します。**Mastercard**は、AIエージェントによる商品探索・購買・決済を支えるAgent Pay/Agent Suiteを通じ、エージェントの登録・認証、権限・利用上限、トークン化などを組み合わせ、**Agentic Commerceの基盤**を整備しています。[*7]

これらの事例は、AIが**組織内のデジタルワーカーから、複数のデジタルワーカーが連携するデジタルワークフォースへ、さらに企業・組織の境界を越えて協働するAgentic Ecosystemへと拡張していく方向性**を示しています。金融業におけるAI活用は、単なる業務効率化から、**AIを新たな「労働力」として組織・企業間の価値創造に組み込む段階**へ移行しつつあります。

*3 Hugh Son（February 6, 2026）“Goldman Sachs taps Anthropic’s Claude to automate accounting, compliance roles”; Srinivas Kumar “Goldman Sachs is quietly moving from using AI as a tool to treating it like a co-worker.”
*4 日本経済新聞（2026年1月26日）「三菱UFJ、同僚は「AI行員」　スピーチライターなど20業務で導入」
*5 babul-prasad（July 15, 2025）“BNY Mellon Leads the Way: The Rise of Credentialed AI Digital Employees in Banking”; CNBC（February 9, 2026）“Digital employees, AI bootcamps: America’s oldest bank is spending billions on tech”
*6 CNBC（June 3, 2026）“Morgan Stanley will soon open its trillion-dollar wealth management funnel to AI agents”
*7 Mastercard（April 29, 2025）“Mastercard unveils Agent Pay, pioneering agentic payments technology to power commerce in the age of AI”; Mastercard（September 10, 2025）“Mastercard unveils new tools and collaborations to power smarter, safer agentic commerce”

3. AIを「使う」から「経営する」へ
―金融業における実践：AI価値の最大化とトークン（Token）経営

金融業では、AIエージェントの活用が実験段階を超え、組織横断の業務遂行や意思決定を支える本格運用へ移行しつつあります。一方、AI投資の効果が間接的・長期的になるほどROIの定量化は難しくなります。KPMGの調査でも、AI投資の効果測定を課題とする組織が前四半期比で32ポイント増加しています。[*8] こうしたなか、投資家や取締役会への価値証明が求められる一方、AIのコストと経済価値を理解し、適切なKPIを設計・測定する「AI Fluency」が、経営層だけでなく組織全体に求められています。

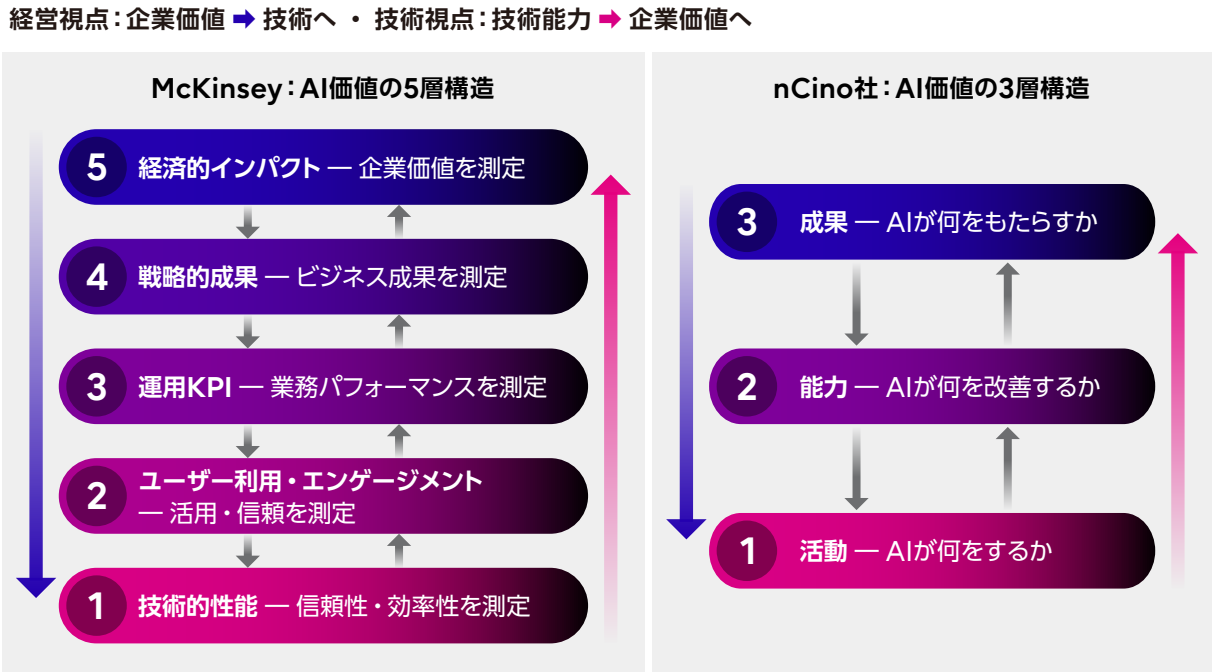
その一方で、AI投資の成果を体系的に測定する先行事例も増えています。そこから得られる重要な示唆は、AIの価値を単一のKPIで捉えるのではなく、**業務・顧客・事業・財務をつなぎ、最終的な経済価値まで追跡すること**です。

本章では、こうした価値測定の先行事例を踏まえ、AIの価値をいかに可視化・最大化し、さらにAIそのものを経営資源として管理するかを考察します。

AI投資に伴う成果測定フレームワークの提案事例：McKinseyとnCino社

Harvard Business Review、McKinsey、nCino社などは、AI投資のビジネスインパクトを測定するためのさまざまなフレームワークを提案しています（図1を参考）。

図1　McKinsey（5層構造）とnCino社（3層構造）の成果測定フレームワーク



出所：McKinseyおよびnCino社のフレームワークをもとに、著者の視点で再構成・作成

*8　KPMG（April 2026）“AI Quarterly Pulse Survey Banking Q1 2026”

McKinseyは、AIの技術的パフォーマンスから財務的インパクトまでをつなぐ5層の「AI Value Stack」を提示しています。[*9]

⑤ **経済的インパクト**：収益増加、コスト削減、利益率、総所有コストなど、AIが生み出す経済価値
④ **戦略的成果**：NPS、顧客満足度、顧客維持率など、事業・顧客成果への影響
③ **運用KPI**：サイクルタイム、欠陥・再作業率、初回解決率、取引当たりコストなど、業務プロセスの改善
② **ユーザー利用・エンゲージメント**：利用率、ワークフローへの浸透率、AIの受容・上書き率など、現場での定着度
① **技術的性能**：誤動作率、レイテンシ、トークンコスト、出力品質など、AIシステム自体の性能

重要なのは、AI投資を導入率や時間削減、生産性向上だけで評価するのではなく、①技術的性能から⑤財務的インパクトまでを一貫してつなぎ、P/Lへの貢献を追跡することです。図1に示すように、経営の視点では⑤財務的インパクトから①技術的要因へ分解し、技術の視点では①から⑤へ価値を積み上げます。すなわち、AI価値の実現にはトップダウンの価値設計とボトムアップの価値実証の双方が必要となります。

一方、金融業向けクラウド銀行プラットフォームを提供する**nCino社**の「銀行業界のAI活用実態調査」では、71%の銀行がAIに関する明確で一貫した評価指標を策定している一方、ROIに直結するKPIの優先度は21%にとどまっています。[*10]

こうした実態を踏まえ、nCino社はAIの価値測定をより簡素化し、**①活動、②能力、③成果**の3層で捉えるフレームワークを提案しています。[*11]

③ **成果**：AIが何をもたらすか。クロスセル、コスト削減、収益増加など、最終的なビジネス価値
② **能力**：AIが何を改善するか。顧客体験、不正検出など、既存プロセスの改善
① **活動**：AIが何をするか。導入・利用状況、タスク完了、従業員生産性、ワークフロー自動化など

両者から浮かび上がるのは、AI投資における「**価値の断絶**」です。すなわち、AIを実装する技術部門と、最終的な事業成果を生み出す事業部門の間で、AIの価値を継続的に追跡・評価する仕組みが十分に確立されていません。

取締役会が求める「**金額換算された財務リターン**」を継続的に説明するには、部門横断で成果を追跡できるデータ基盤と、共通の評価基準・プロセスを、AI導入と同時に設計することが不可欠です。

AI価値測定フェーズ1の事例：JPMorgan / Lloyds

AIの活動から財務的価値までを体系的に測定する先行事例も現れています。

JPMorgan（米国）は、600超の実稼働AIユースケースにビジネス成果指標を設定し、**①時間節約、②コスト削減、③収益貢献、④価値創出額**の4カテゴリーで評価しています。個別ユースケースの価値を積み上げ、AIによる価値創出額は約20億ドルと、年間AI投資額に近い水準に達したと報告しています。[*12] これは、成果をボトムアップで積み上げROIを検証する**「財務規律型」アプローチ**の典型例です。

Lloyds Banking Group（英国）は、AI活用を**①意識・関与、②習熟・定着、③ビジネス価値・顧客インパクト**の3段階で追跡し、利用率だけでなく、業務への統合、顧客満足度、新サービス創出、時間・コスト削減、利益貢献まで測定しています。2025年には50超の生成AIユースケースから約5,000万ポンドの価値を測定し、2026年には生成AIとAgentic AIによる価値が1億ポンドに達すると見込んでいます。[*13]

*9 McKinsey（April 24, 2026）“From promise to impact: How companies can measure—and realize—the full value of AI”
*10 nCino（2026年5月13日）「銀行業界のAI活用：銀行の経営陣が語る今後の展望」
*11 nCino（August 20, 2026）“From AI Adoption to AI Revenue: The Three Layers of Banking KPIs”
*12 Paul Brzozowski（February 11, 2026）“JPMorgan Spent $2B on AI. Here's What They Measured.”
*13 Trystan Davies and David Blott（April 7, 2026）“How is AI changing the future of work in banking?”

両社のアプローチには違いがあります。JPMorganは生産性・収益への直接的な価値を重視し、LloydsはAI活用から顧客・事業・財務成果までを広く追跡しています。しかし**共通するのは、個別ユースケースの成果をボトムアップで積み上げ、最終的な財務価値へ接続すること**です。これは、AI投資の価値を**「測定し、証明する」Phase 1**と位置づけられます。

しかし、AI活用が組織全体へ拡大すると、次の問いは「どれだけ価値を生んだか」から**「限られたAI資源をどこに配分し、いかに価値を最大化するか」**へ移ります。これが、**Phase 2としての「Token経営」**です。

金融業にとってのトークン経営：価値創出の最大化

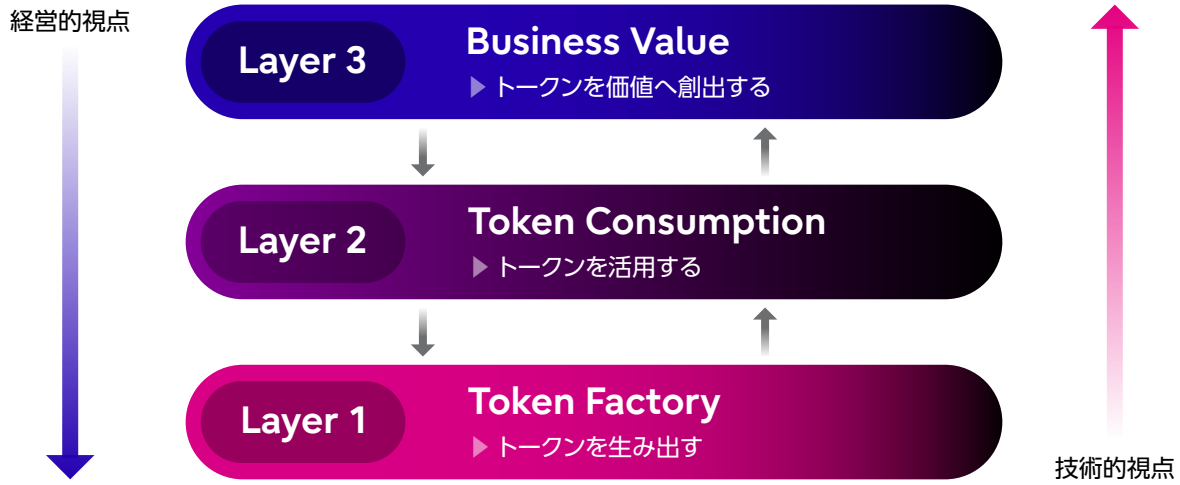
AIエージェントが実験段階から企業全体へ広がるなか、金融業もAIを「使う」段階から「経営する」段階へ移行しつつあります。その鍵となるのが、AIの計算・利用を支えるトークンを新たな経営資源として捉える視点です。AIエージェントが自律的な「デジタル労働力」として拡大するにつれ、トークンのコストと価値を管理することは、経営・CFOが担う新たな経営規律となります。[*14]

(1) AIの価値創造を支える3層構造とトークンの役割

トークンは、単なるAI利用量を示す技術的単位ではありません。**AIリソースの生成・配分・活用から事業価値への転換までをつなぐ共通の経営指標**となり得ます。重要なのは、AIを大量に使うことでも、トークンコストを削減することでもなく、**限られたAIリソースをいかに効率的に生み出し、最適に配分し、最大の事業価値へ転換するか**です。[*15]

図2は、この考え方を①**Token Factory**、②**Token Consumption**、③**Business Value**の3層構造として示したものです。

図2　AIの価値創造を支える3層構造とトークンの役割



出所：著者作成

① **Token Factory**：トークンをいかに効率的に生成するか
② **Token Consumption**：トークンをいかに最適に配分・活用するか
③ **Business Value**：AI利用をいかに事業価値へ転換するか

*14 Muthumari S & Natasha Nayak (August 6, 2026)
"Tokenomics for Banking AI: Why Banks Need a New Economic Discipline"
*15 富士通（2026年9月4日）「AIエージェント時代のトークン経営：企業競争力を左右する新しいマネジメント」

(2)トークン経営の主要KPIと管理ポイント

表3は、各レイヤーの主要KPIと管理ポイントを整理したものです。**技術KPIと経営KPIをつなぎ、各レイヤーで何を最適化するかを可視化すること**が、トークン経営の実践基盤となります。

表3　トークン経営を支える主要KPI

レイヤー	主なKPI	管理のポイント
Layer 3 **Business Value** **(事業価値)**	① トークンコスト当たり売上 (Revenue per Token Cost) ② 生産性向上率(Productivity Gain) ③ 品質改善指標(Quality Improvement) (例:初回解決率、コード採用率)	**トークンを最大の事業価値へ転換する** ・ROIの最大化 ・顧客体験・従業員体験の向上 ・事業・部門別AI投資効果の可視化
Layer 2 **Token Consumption** **(トークン利用)**	① 1タスク当たり平均トークン数 (Average Tokens per Task) ② キャッシュヒット率(Cache Hit Rate) ③ トークン再利用率(Token Reuse Rate)	**必要最小限のトークンで最大の業務成果を実現する** ・キャッシュ最適化 ・ルーティング最適化 ・不要な推論・再計算の削減
Layer 1 **Token Factory** **(トークン生成基盤)**	① 100万トークン当たり総コスト (Cost per Million Tokens) ② 初回トークン出力時間 (Time to First Token:TTFT) ③ GPU稼働率・スループット (GPU Utilization & Throughput)	**トークンを低コスト・低遅延・高効率で供給する** ・コンピューティング資源の最適化 ・推論性能の向上 ・安定したトークン供給基盤の構築

出所:著者作成

(3)金融業におけるトークン経営の先行事例

トークンをAI活用の経営資源・コスト指標として管理する先行事例も現れています。

JPMorganは、独自の**AI Traffic Control Layer**を構築し、モデルルーティング、セマンティック・キャッシュ、ガバナンスを一元管理しています。[*16] これにより、用途に応じたモデル選択、AI利用・コストの可視化、トークン消費の最適化を進めています。さらに、トークン消費だけでなく、時間削減や経済価値まで追跡し、AI支出を成果と結びつける**財務規律**を構築しています。JPMorganは、Tokenを「**配分・最適化すべき経営資源**」として管理する先行例といえます。

中国平安保険(Ping An)は、「AI in ALL」戦略のもと、**Token Factory→Token Consumption→Business Value**を一体で管理しています。基盤モデルや推論、コンピューティングを最適化するとともに、顧客サービス、AI診療、不正検出などへAIを展開。処理時間を50%以上短縮し、不正検出では**71.1億元の保険金削減**を実現しています。[*17] さらに、コンピューティング能力単位当たりのトークン出力を**368%増加**させる一方、トークン当たりコストを**76%削減**しています。Ping Anは、Tokenの生成・消費を最適化し、**AIリソースを事業価値へ転換する**先行例といえます。

*16 Shikhar Gupta (July 14, 2026)
"JPMorgan expects industry to manage AI spend with more 'discipline' as it eyes Singapore tech talent"; Dinesh Elumalai (September 1, 2026)
"The AI Gateway Pattern That JPMorgan, Stripe, and Every Smart Fintech Is Quietly Standardizing On"
*17 Ping An (September 23, 2026) "Ping An CTO Ray Wang: From Tokens to Value Creation"

4. AI時代のガバナンス：金融業の新たな競争力
―リスク対応から「AIレジリエンス」へ

世界最大のホワイトハッカーコミュニティを持つHackerOneは、2025年10月の「Hacker-Powered Security Report」で、AIの自律性が高まるなか、HackerOneのプラットフォーム上で報告されたAI関連脆弱性が210%増加したと報告しています。[*18] AIモデルの高度化とグローバルなフロンティア競争が進むなか、ガードレールの回避、間接プロンプトインジェクション、外部ツールやシステムの不正操作など、新たなリスクも顕在化しつつあります。

ガバナンスや規制の整備を待つのではなく、企業自らが先手を打って対応することが重要です。AIの高度化は、金融機関の競争力を高める一方、新たなリスクの源泉にもなり得るという二面性を持つからです。

AIエージェント時代におけるリスクと被害の連鎖・増幅

金融機関にとって、AIエージェント時代のリスクは、従来のサイバー攻撃の延長ではなく、**速度・規模・対象・被害の形そのものを変える**可能性があります。

AIエージェントは、権限の逸脱、誤判断の自動拡散、データ漏えい、エージェント間の連鎖事故などを通じて、**「最大の生産性レバー」であると同時に、新たな内部統制リスク**にもなります。さらに、APIキー、トークン、SSOなどの認証情報を保持し、複数のSaaSや社内システムにアクセスするエージェントが侵害されれば、権限の悪用や横展開の起点となり得ます。単一の侵害が、エージェント間連携や異なる環境への拡散を通じて、**連鎖的に被害を増幅するリスク**への対応が必要です。

金融AIセキュリティの新パラダイム：信頼の再定義

こうした変化は、金融業におけるAIの「Trust」の考え方そのものを変えつつあります。

従来、組織内部の誤作動を防ぐ**Safety**と、外部からの攻撃を防ぐ**Security**は別々に扱われてきました。しかし、AIの不完全性を突く高度な攻撃や自律的な攻撃手法の登場により、その境界は曖昧になっています。ゼロトラストの考え方においても、両者は**「誰を、何を、どこまで信頼するか」というガバナンスとTrustの課題**として統合的に捉える必要があります。

また、金融規制において100%の防御を前提とすることは現実的ではありません。重要なのは、事前にリスクを抑える**Risk Management**に加え、重大なインシデントが発生しても業務を継続し、迅速に復旧する**Digital Operational Resilience**を一体的に構築することです。

さらに最前線では、自律型AIを悪用する「**AIで武装した攻撃者**」と、人間の監督のもとAIで防御能力を高める「**AIで拡張された防御者**」によるリアルタイムの攻防が始まっています。こうした状況では、**Trust・Security・Resilienceを統合した設計**が不可欠です。具体的には、①攻撃の兆候や意図を早期に捉える**事前の備え**、②侵入や被害拡大を阻止する**事中の防御**、③侵害後の封じ込めと復旧を担う**事後対応**を、一連の仕組みとして設計・運用する必要があります。特に、自律型AIによる攻撃に対しては、**AIを前提とした防御能力の高度化**が求められます。

AIセキュリティの新たな課題に対応するため、先行企業は実践的な防衛モデルやフレームワークの構築を進めています。

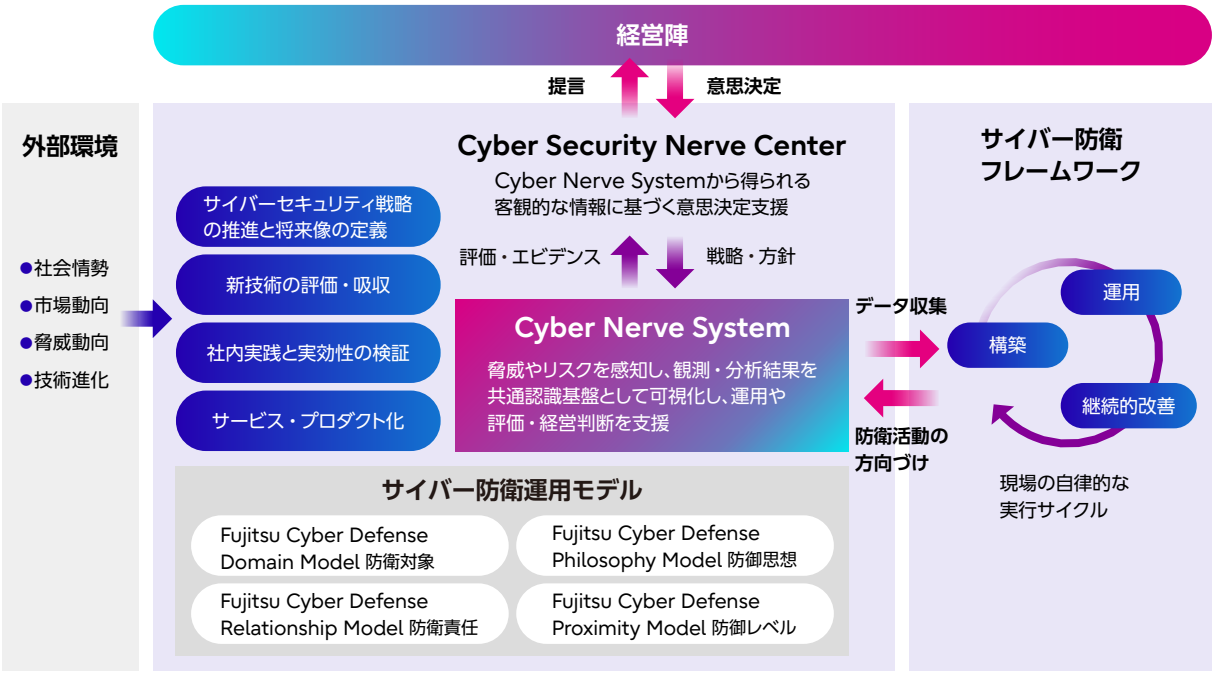
例えば、富士通は、生成AIの進化や企業システムの複雑化を踏まえ、サイバー防衛能力を継続的に進化させる戦略を策定しました。[*19] 特徴は、セキュリティをシステム構築後の対策に限定せず、**戦略立案から設計・開発、実装、運用、評価・改善までを一体化した「サイバー防衛ライフサイクル」**として捉える点です。

*18 HackerOne（October 1, 2025）
"HackerOne Report Finds 210% Spike in AI Vulnerability Reports Amid Rise of AI Autonomy"

*19 富士通（2026年8月26日）「AI時代の企業のサイバー防衛ライフサイクルを実現するサイバーセキュリティ戦略を策定 Cyber Security Nerve Centerを核に、富士通の実践知をお客様へ展開」

さらに、全社的な防衛能力の構築・運用・改善を体系化した**サイバー防衛運用モデル**と、その実践を支える**サイバー防衛フレームワーク**を定義しました（図3を参照）。加えて、得られたデータから脅威やリスクを検知し、分析結果を可視化して防衛活動を方向づける仕組みとして、現場の自律的な実行・改善を促す**Cyber Nerve System**を構築しています。

図3　富士通が策定したサイバーセキュリティ戦略の全体像



出所：富士通プレスリリース（2026年8月26日）

金融AIの社会実装を支える多層的ガバナンス

こうしたリスクに対応するため、金融業におけるAIエージェントの社会実装では、自律性や利便性の追求と同時に、**制御可能性（Security）**、**監査可能性（Compliance）**、**説明責任（Trust）**を組み込むことが重要です。[*20]

特にAIエージェントが金融取引や顧客サービスに直接関与するほど、以下のような多層的なガバナンス基盤が求められます。

Identity Management：M2M/AI2AI認証などにより、人間とAIエージェントのID・権限を明確に分離
Authorization Management：
Pre-transaction Policy Enforcementにより、取引実行前に権限・取引条件・承認要件を検証
Auditability：Tamper-evident Audit Recordsにより、AIエージェントの行動履歴や意思決定を記録・追跡
Trust Infrastructure：Zero TrustやKYAなどを組み込み、AIエージェントの認証・信頼性を管理

さらに、AIを活用してセキュリティ能力を高める「**バイオニックハッカー**」（AIによってサイバー能力を拡張された次世代のホワイトハッカー）など、外部の専門知識を、バグバウンティ（脆弱性報奨金制度）などを通じて制度的に取り込むことも、AI時代の防御力を高める有効なアプローチとなります。

*20 Justin Kirsch（July 30, 2026）“The AI Agent Identity Crisis: Machine IDs Will Outnumber Humans by 2027”;
JPMorgan Chase（March 23, 2026）“Securing the next generation of AI agents”

おわりに
AI経営に向けて―金融業リーダーへの3つの提言

本稿では、AIを組織に組み込み、デジタル労働力として活用し、経営資源として管理するとともに、AIレジリエンスを持続的な競争力へとつなぐ方向性を考察してきました。これを踏まえ、金融業リーダーへの提言を3つ示します。

(1) 人間とAIの協働によるハイブリッド型組織の構築

AIファースト時代の価値創出は、人間とAIの代替ではなく、双方の強みを生かす**補完的な協働**によって最大化されます。そのためには、組織・ワークフローを再設計し、人間とAIそれぞれの役割、責任、権限、KPI、意思決定範囲を明確にしたハイブリッド型モデルの構築が不可欠です。

その実現には、経営層のリーダーシップと組織全体の意識変革に加え、**試行・検証・フィードバック・改善・スケールアップ**のサイクルを迅速に回すことが重要です。

(2) AI価値の測定・証明と資源配分の仕組みの確立

AI活用が本格展開へ移行するなか、投資効果を定量化し、財務価値への貢献を証明することが求められています。その基盤として、部門横断で成果を追跡できるデータ基盤と、共通の評価基準・プロセスをAI導入と同時に整備する必要があります。

実践に当たっては、まず個別ユースケースの価値を**「測定・証明」するPhase 1**から始め、その知見をもとに、AIリソースを**「配分・最適化」するPhase 2**へ進化させます。さらに、測定結果をダッシュボードなどで可視化し、経営層から現場までが日常的な意思決定と改善に活用することが重要です。

(3) AIレジリエンスを支えるガバナンスの強化

高度な自律型AIは金融機関の競争力を高める一方、新たなリスクの源泉にもなります。AIによる競争優位を持続させるには、**事前の備え・事中の防御・事後の対応と復旧**を一体化し、AI時代に適した防御能力とレジリエンスを構築する必要があります。

その基盤として、制御可能性・監査可能性・説明責任を組み込み、人間とAIのID・権限管理、取引前のポリシー検証、行動・意思決定の記録と追跡、エージェントの認証・信頼性管理を一体的に運用することが不可欠です。さらに、バグバウンティなどを通じて外部の専門知識を取り込み、AIで拡張された防御態勢を強化することも求められます。

AIを経営することの先に、AIとともに経営する時代が到来するかもしれません。その実現可能性は、AIが生み出す価値の最大化と、それを支えるガバナンスをいかに成功させるかにかかっています。

著者紹介



金 堅敏(Jianmin Jin)博士

2020年～　　富士通株式会社 チーフデジタルエコノミスト
1998年～2020年　富士通総研 主席研究員

理工系と社会科学系の知見を融合する金 堅敏博士は、グローバルかつビジネスの視点からソートリーダーシップ活動を展開しています。グローバル経済、デジタルエコノミー、デジタルイノベーション、そしてAI時代の企業トランスフォーメーションを主なテーマに、リサーチと深いインサイトをもとに、インサイトペーパーの執筆、メディア寄稿、講演、コンサルティングなどを通じて、ビジネスリーダーに新たな視点と実践的な示唆を発信しています。

直近の著作物：以下の富士通ホワイトペーパー、ほか。

- 次世代AI産業競争力の構築 ― 米中AI競争から読み解く新たな競争原理（2026年10月）
- AIエージェント時代のトークン経営：企業競争力を左右する新しいマネジメント（2026年9月）
- AIが変える半導体市場と企業の競争戦略（2026年7月）
- AI時代の創薬2.0：AIネイティブ創薬企業が再構築する次世代創薬モデルと競争優位（2026年7月）
- 製造業におけるAI労働力の進化（2026年6月）

著者は、このインサイトペーパーの作成中に貴重な助言や日頃の揺るぎないご支援をしてくださった山本 義規、水野 義博、佐藤 丈師、猪瀬 森主、新田 隆司、目黒 紘子、佐藤 由起子に感謝申し上げます。

FUJITSU

記載されている企業名・製品名などの固有名詞は、各社の商標または登録商標です。
本資料は発行日現在のものであり、富士通によって予告なく変更されることがあります。
本資料は情報提供のみを目的として提供されたものであり、富士通はその使用に関する責任を負いません。
本資料の一部または全部を許可なく複写、複製、転載することを禁じます。
富士通および富士通ロゴは、富士通株式会社の商標です。

©Fujitsu Limited　2026年11月 v1.0