2026 年 4 月 15 日　全 8 頁

金融分野における AI 規制の在り方

「AI ディスカッションペーパー（第 1.1 版）」の要点と国際比較

金融調査部　研究員　谷 京

[要約]

- 生成 AI の技術的進化と金融分野への浸透が急速に進み、顧客向けサービスへの生成 AI 利用についても、範囲・条件を絞ったサービス提供やその検討が行われる段階に至っている。このような環境の変化を踏まえ、金融庁は 2026 年 3 月 3 日、前年 3 月に公表していた「AI ディスカッションペーパー」を第 1.1 版に改訂した。

- AI ディスカッションペーパーでは、顧客向けサービスを念頭に置いたリスク低減策や、AI 活用に関して金融機関から寄せられた法令解釈上の疑問点に対する金融庁の見解が示された。また、データマネジメントの重要性や AI エージェントの今後の展望についても言及された。

- AI の金融分野への適用をめぐる規制・ガイドライン整備は、国際的にも進展している。英国では、金融分野における AI 利活用を促進する観点から、既存の規制枠組み内でのプリンシプルベース・アプローチによる柔軟な対応が図られている。これに対し、米国は AI 規制の新設を避けながらも、共通実務標準の策定による監督・検査基準の実質的導入を進めている。さらに、EU では包括的なハードロー（人工知能法）による事前規制で、金融分野での AI 利用に際しての要求事項が明確化されている。

- AI ディスカッションペーパーは、生成 AI の顧客向けサービスへの展開や AI エージェントの台頭といった急速な環境変化を踏まえ、金融機関の積極的な AI 活用を後押しする姿勢が一貫している。ただし、AI がもたらすリスクの具体的な定義や分類、金融機関の法的義務については、米国のような共通実務標準の策定による明確化が今後の課題として残されていよう。

株式会社大和総研 丸の内オフィス　〒100-6756 東京都千代田区丸の内一丁目 9 番 1 号 グラントウキョウ ノースタワー

1. 金融庁「AI ディスカッションペーパー」改訂の経緯

近時における生成 AI の急速な普及を受け、金融庁は金融分野における健全な AI 活用の促進に向けた取組みを本格化させている。2024 年 10 月から 11 月にかけて金融機関へのアンケート調査を実施し、2025 年 3 月 4 日にはその結果を踏まえた「AI ディスカッションペーパー（第 1.0 版）」を公表した。同文書は、「AI が、将来的に金融サービスの提供の在り方や金融機関等のビジネスモデルを抜本的に変革しうる技術であるとすれば、〔中略〕技術革新に取り残されて中長期的に良質な金融サービスの提供が困難になる『チャレンジしないリスク』も十分に認識されるべきである」（4 ページ）ことを前提に、金融機関が規制やリスクへの懸念から過度に萎縮することなく、健全な AI 活用を進めるためのセーフハーバーを提供することを主眼としていた[1]。

もっとも、上記のアンケート調査によれば、当時すでに生成 AI の導入は進展していたものの、その用途の大半は文書の作成・要約や翻訳、社内 FAQ 等の情報検索といった「社内利用（業務効率化等）」に留まっていた。一部の金融機関はコールセンター業務支援や稟議書等の作成といった「対顧客サービスへの間接的な利用」にも取り組んでいたが、顧客に対して直接的に AI によるサービスを提供する事例は極めて限定的であった[2]。

しかしながら、生成 AI の技術的進化と金融分野への浸透は、当初の想定を上回る速度で進行した。すなわち、2025 年 6 月から 12 月にかけて開催された「金融庁 AI 官民フォーラム」では、2024 年のアンケート調査時点ではほとんど行われていなかった顧客向けサービスへの生成 AI 利用について、範囲・条件を絞ったサービス提供またはその検討が行われる段階に至っていることが明らかとなった。もはや生成 AI は単なるテキスト処理を超え、一部の顧客対応や営業支援も可能なツールへと進化を遂げており、将来的には「顧客への資産運用に関する相談対応やコンサルティングなど、より金融サービスの本業に近い領域で生成 AI を直接顧客向けに活用するような高度なサービス」（金融庁「AI ディスカッションペーパー（第 1.1 版）」2026 年 3 月、19 ページ）を提供することも想定されるようになったのである[3]。

このような環境の変化を踏まえ、金融庁は 2026 年 3 月 3 日、AI 官民フォーラム等で収集した最新の知見や事例を反映し、金融分野における AI 活用の現状と課題をアップデートした「AI ディスカッションペーパー（第 1.1 版）」を公表した。第 1.1 版というナンバリングからも看取されるように、同文書は第 1.0 版に引き続き、「中長期的に金融庁の政策や金融機関等における AI ガバナンスの在り方を検討する上での土台」（7 ページ）として位置付けられている[4]。

そこで、本稿では第 1.0 版から追加・更新された主要な論点を紹介するとともに、証券監督者国際機構（IOSCO）や外国規制当局のガイドラインとの比較を行い、金融分野における AI 規制の在り方を考察する。

---

[1] 金融庁「AI ディスカッションペーパー（第 1.0 版）」2025 年 3 月、4-5 ページ。
[2] 同上、17-19 ページ。
[3] 金融庁「AI ディスカッションペーパー（第 1.1 版）」2026 年 3 月、7・18-19 ページ。
[4] 同上、7 ページ。

2. 第 1.0 版から追加・更新された内容

（1）生成 AI の顧客向けサービスへの活用に関する指針の強化

第 1.1 版では、金融機関が顧客向けサービスに生成 AI を活用する際の留意点が追記された[5]。ここで注目すべきは、顧客向けサービスを念頭に置いたリスク低減策を、AI のライフサイクルに沿って次の四つの局面に分解し、それぞれの段階における具体的な検討事項と対応策を整理した点にある。

第一の局面は、AI システムの設計と事前のテスト・検証である。ここでは、生成AI特有の不確実性やハルシネーション（一見もっともらしいが事実に基づかない偽情報の生成）、不適切な出力リスクを構造的に排除するための技術的アプローチが明記された。具体的には、回答の範囲を内部データ等で厳格に制限する RAG（検索拡張生成）技術の活用や、より高度な LLM（大規模言語モデル）の選択、不適切・断定的な回答のフィルタリングといった重層的なガードレール（安全策）を設計することが推奨されている。また、特にサービス導入初期には、顧客が AI 対応と人手対応を選択できる体制を整えつつ、低リスク業務から中核業務へ段階的に生成 AI 活用を拡大していくリスクベースのアプローチも重要視されている。

第二の局面は、顧客への適切な説明・注意喚起である。ここでは、AI が生成した回答であることを顧客に対して明示し、その出力に誤りや最新でない情報が含まれる可能性について顧客へ注意喚起することが求められている。また、対話の中で顧客の理解度を確認するステップを設け、必要または顧客の希望に応じて人間の担当者へ切り替えられる仕組みを用意することや、回答の根拠となった情報源や参照データを顧客側から確認できる機能を実装することも盛り込まれた。

第三の局面は、検証・モニタリングである。AI モデルは運用環境におけるデータの変化によって、性能が劣化したり予期せぬ挙動を示したりする性質を持つ。そのため、サービス提供後においても、AI の振る舞いを継続的に監視する仕組みの構築が要求される。具体的には、AI と顧客の対話ログを保存・モニタリングし、例えば特定の金融商品を過度に推奨していないか、推奨ロジックの文書化や第三者レビューによるチェックを行うことが提案されている。

第四の局面は、これら一連の開発・運用プロセス全体を俯瞰するガバナンスである。AI の本格活用にあたっては、経営陣を含む全社的な体制整備と現場職員のリテラシー向上が不可欠であることが強調されている。また、アジャイルなガバナンス、つまり環境変化に応じてリスク管理策や内部ルールを迅速・柔軟に更新できる統治体制の必要性が指摘され、AI のライフサイクルを通じたリスクマネジメントも求められている。

---

[5] 金融庁「AI ディスカッションペーパー（第 1.1 版）」2026 年 3 月、37-38 ページ。

（2）法令解釈の明確化

第 1.1 版では、AI 活用に関し、金融機関から寄せられた法令解釈上の疑問点に対する金融庁の見解が追加された[6]。例えば、証券会社等の金融機関が自社のシステム関連子会社等に AI モデルの開発や機械学習を委託する際、非公開情報を含む機微な顧客データを提供することは非公開情報の授受規制（金融商品取引業等に関する内閣府令第 153 条第 1 項第 7 号）に抵触するのではないかという懸念がある。この点、第 1.1 版では、一定の厳格なガバナンス要件や情報アクセス制御、監査体制を満たすことを前提に、グループ内のシステム子会社への顧客データの提供を許容する方向性が明記された。

また、生成 AI が直接顧客に金融商品を推奨する場合、それが金融商品取引法上の「勧誘」に当たるかという問いに対しては、「個別具体的な事案に即して実質的に判断する」（40 ページ）との見解が提示された。金融庁は、AI が投資家の判断に与える影響や AI への営業担当者等の関与の度合いなどを考慮して、勧誘への該当性を判断するとのスタンスを示したといえる。

（3）データマネジメントの重要性

金融機関が生成 AI の活用に際し、外部事業者が提供する汎用 LLM をそのまま利用する場合、社内データは基本的に不要である。しかし、自社の業務プロセスに最適化するための RAG やファインチューニング、コンテキスト内学習等を試みる場合は、AI 活用に適したデータベースの構築と十分な学習データの確保が必要となる。

その中で、不正確な情報やバイアスを含んだデータからは、どれほど高度な LLM を用いても適切な出力は得られない（ガベージイン・ガベージアウト）。したがって、AI 活用の前提として、明確なビジネス目的を持った体系的なデータ整備が求められる。このように、第 1.1 版では、AI モデルの有用性と出力の正確性は学習データの質に依存するという AI の技術的特性を踏まえ、データマネジメントの重要性が強調された[7]。

（4）AI エージェントへの言及

第 1.1 版では、2025 年が「AI エージェント元年」となったことを受け、AI エージェントに関する節が新設された[8]。ここでは、AI エージェントを「特定の目標を達成するために自律的に行動する AI システム」と定義し、LLM が AI エージェントに組み込まれることで幅広い業務に応用できるとしている。

さらに、複数の AI エージェントが互いに連携しながら自律的に処理を繰り返す「エージェンティック AI」が実現すれば、金融分野における多くの業務プロセスを自動化できる可能性があ

---

[6] 金融庁「AI ディスカッションペーパー（第 1.1 版）」2026 年 3 月、39-41 ページ。
[7] 同上、23-24 ページ。
[8] 同上、19 ページ。

るとの期待が示された。今後は、こうしたAIエージェントやエージェンティックAIを顧客向けサービスへと広げていく動きが見込まれるほか、人手不足への対応という観点からも、業務への本格的な導入を検討していく必要があると指摘されている。

3. IOSCOおよび外国規制当局の関連規制・ガイドラインとの比較

AIの金融分野への適用をめぐる規制・ガイドライン整備は、国際的にも急速に進展している。以下では、IOSCOや外国規制当局のアプローチを整理し、金融庁のAIディスカッションペーパーとの共通点および相違点を抽出する。

(1) IOSCO：ガバナンスと経営陣の説明責任

IOSCOは2021年、「人工知能及び機械学習を利用する市場仲介者及び資産運用会社向けのガイダンス」を公表し、金融仲介業者および資産運用会社に対しAI・機械学習に関するベストプラクティスを提示した。ここでは、①経営陣による明確な監督責任の確立、②AI・機械学習の十分なテストおよび継続的な検証、③十分なスキル・専門性・経験の確保、④第三者提供者への依存関係の適切な管理、⑤顧客・投資者への適切な情報開示、⑥データ品質とバイアスの管理という六つの行動規範が定められた。さらに、2025年には同ガイダンスを土台とする市中協議文書「資本市場における人工知能：ユースケース、リスク及び課題」が公表された。同文書の主目的は、生成AIやLLMといった新たなAI技術が投資者保護、市場の健全性、および金融システムの安定性にもたらす潜在的リスクについて、加盟当局間で共通の理解と監視水準を形成することにある。

IOSCOのガイドラインは、AIの高度化・不透明性を前提としつつも、経営陣の直接的な説明責任を重く見ている点に特徴がある。具体的には、AI・機械学習の設計、テスト、導入、監視、および統制に対する監督責任を負う「上級経営陣」を指定し、アルゴリズムがもたらす結果を継続的に検証（バリデーション）することを強く求めている。また、コンプライアンス部門とリスク管理部門がアルゴリズムを理解し、これを検証・牽制できる体制を整えるべきであるとしている[9]。

金融庁のAIディスカッションペーパーが経営陣を含む全社的なガバナンスの構築を求めている点は、IOSCOの行動規範と軌を一にしている。また、上記④の第三者提供者への依存関係の適切な管理は、金融庁が示したシステム関連子会社等への委託に関する法令解釈と同様の問題を扱っている。両者ともに、基盤モデルを提供するテック企業への依存やAIモデルの不透明性が増したとしても、金融機関は説明責任を負っているという立場を取っている。

---

[9] IOSCO, “The use of artificial intelligence and machine learning by market intermediaries and asset managers: Final Report,” 2021/9, pp. 17-20.

（2）EU：包括的ハードロー

欧州連合（EU）は 2024 年 8 月、世界初となる包括的かつ法的拘束力を持つ人工知能法（AI Act）を発効させた[10]。同法はリスクベース・アプローチをすべての産業に対して横断的に適用し、AI システムを「許容できないリスク」「高リスク」「限定的なリスク」「最小限のリスク」の 4 段階に分類している。

個人の信用評価（クレジットスコアリング）や生命保険・健康保険の引受リスク評価などに用いられる金融 AI は、高リスクに分類される可能性が高い。高リスクの AI システムに対しては、リスクマネジメントシステムの導入・運用、データガバナンス、技術文書の作成と当局等への提出、適切な透明性確保、人間による監視などが求められる。これらの規制は、原則として 2026 年 8 月 2 日から適用が開始される。

EU の人工知能法で整理された AI システムのリスク、すなわちアルゴリズムのバイアス、データ品質、出力の不透明性、プライバシー／セキュリティ等の課題は、金融庁の AI ディスカッションペーパーにも共通する懸念事項である。しかし、金融庁が既存業法の枠組みで個別具体的な事案に即して AI リスクを吸収・管理する「技術中立」の姿勢を取っている一方で、EU は包括的なハードローによって高リスク AI への厳格な要求事項を課している点は、監督手法上の大きな相違である。

（3）FCA：プリンシプルベース・アプローチとイノベーション促進

英国の金融行為規制機構（FCA）は、企業が技術の変化や市場の発展に適応できる柔軟性を提供するため、AI に特化した追加規制を設けないとの姿勢を明確に打ち出している。代わりに、AI がもたらすリスクに対しては、消費者義務（Consumer Duty）やシニアマネージャーおよび認証レジーム（SM&CR）といった既存の規制枠組みを適用し、対処する[11]。

また、FCA は 2025 年 4 月、「AI ライブテスト（AI Live Testing）」という規制サンドボックス環境を立ち上げた。AI ライブテストは、金融機関等が実際の市場・顧客環境に近い空間で AI ソリューションを試験運用できる仕組みであり、規制当局と事業者が対話を通じて AI リスクを共同で検証・軽減していくとされている。AI ライブテストの評価レポートは、2 段階にわたる試験運用の実施結果を踏まえ、2027 年第 1 四半期に公表される予定である[12]。

詳細かつ規範的なルールを策定するのではなく、既存の規制枠組みを AI に適用・拡張していくプリンシプルベース・アプローチにより、金融分野における AI の安全かつ責任ある導入を支援するという点において、金融庁と FCA の方向性は共通しているといえよう。

---

[10] 詳細は、矢田歌菜絵「EU AI 規則（AI Act）公布」大和総研レポート、2024 年 8 月 5 日を参照されたい。
[11] FCA, “AI and the FCA: our approach,” 最終更新日 2026/2/13.
[12] FCA, “FS25/5: AI Live Testing,” 最終更新日 2026/3/16.

（4）米国：共通実務標準のリスク管理フレームワーク

米国では、AI に特化した包括的な規制法は存在せず、各規制当局が所管する領域において既存の規制枠組みを適合させ、AI リスクに対応している。金融分野では 2026 年 2 月、米国財務省の主導の下、連邦準備制度理事会（FRB）や金融機関 100 社超も関与する官民協働の「金融サービス AI リスク管理フレームワーク（FS AI RMF）」が公表された。

FS AI RMF は、米国国立標準技術研究所（NIST）の AI RMF を金融分野向けに拡張したものである。その核心は、AI のライフサイクル全体をカバーする 230 の「コントロール目標」にある。すなわち、AI の利用が金融機関の意思決定や顧客・市場に及ぼす影響を踏まえ、AI の企画・設計から運用・廃止に至るまでの各段階において考慮されるべき統制と説明責任が具体化された。これらのコントロール目標は、AI の利用段階（①導入初期、②限定的活用、③高度化・発展、④業務への組込み・定着）に応じて水準を変更する設計となっており、中小金融機関にも配慮した比例原則が採用されている[13]。

金融庁の AI ディスカッションペーパーと FS AI RMF は、AI のライフサイクルに沿ったリスク管理の重要性を共有している。また、両者ともに法的拘束力は有していない。もっとも、FS AI RMF の 230 項目にも及ぶコントロール目標は、金融庁のプリンシプルベース・アプローチと比較して、より詳細なチェックリストとしての性格を帯びている。換言すれば、FS AI RMF は金融分野における AI リスク管理の共通実務標準を整備したことで、今後の監督・検査における事実上のベンチマークとなる可能性が高い。

4. 今後の展望

このように、金融庁の AI ディスカッションペーパーで提示された論点は、国際的にも主要な論点と合致している。とりわけ AI ガバナンス、バイアスとデータ品質、金融機関の説明責任、透明性および利用者保護といった課題は、各国・地域で共通して重視されている。

他方で、具体的な規制手法には違いもある。日本や英国では、金融分野における AI 利活用を促進する観点から、既存の規制枠組み内でのプリンシプルベース・アプローチによる柔軟な対応が図られている。これに対し、米国は AI 規制の新設を避けながらも、共通実務標準の策定による監督・検査基準の実質的導入を進めている。さらに、EU では包括的なハードロー（人工知能法）による事前規制で、金融分野での AI 利用に際しての要求事項が明確化されている。

その中で、金融庁 AI ディスカッションペーパー（第 1.1 版）は、AI 官民フォーラムという官民対話の成果を反映した。同文書は、生成 AI の顧客向けサービスへの展開や AI エージェントの台頭といった急速な環境変化を踏まえ、「チャレンジしないリスク」（4 ページ）という強いメッセージとともに、金融機関の積極的な AI 活用を後押しする姿勢が一貫している。もっとも、AI がもたらすリスクの具体的な定義や分類、金融機関の法的義務については、詳細かつ規

[13] Cyber Risk Institute, “The CRI Financial Services AI Risk Management Framework,” 2026/2.

範的なルールを策定するものではない。既存の規制枠組みをAIに適用・拡張していくプリンシプルベース・アプローチを継続するとしても、米国のような共通実務標準の策定による明確化が今後の課題として残されていよう。