2024 年 9 月 24 日

公益財団法人 金融情報システムセンター（FISC）

金融機関による AI の業務への利活用に関する安全対策の観点からの考察

はじめに

（１） 金融機関による AI の業務への利活用を巡る動き

AI（Artificial Intelligence、以下、「AI」という。）を事業に利活用する企業が増えている。金融機関においても、調査や実証実験を推進する先が増えていることに加え、実際に業務に AI を活用する動きが進んでいる。

AI に関する動向をやや長い目で振り返ると、まず、2000 年代以降、ディープラーニング等の技術を用いた「画像認識」、「自然言語処理（翻訳等）」、「音声認識」の仕組みが開発され、これらを活用し、予測、提案又は決定を行うことができるという、特定の分野に特化したシステムが、AI と認識されるようになった。その後、2021 年以降、大規模言語モデルに代表される基盤モデルの台頭により、特定の分野のみに特化した AI ではない、汎用的な AI の開発が進んでいる。こうしたなかで、画像、文章等を生成する「生成 AI」が普及するようになり、生成されたアウトプットの質の高さと、新しいモデル及びツールが投入されるスピードの速さとにおいて、世界的な注目を集めている。実際、最近、提供されるモデルやサービスが多様化し、選択肢が拡がっている。

わが国の金融機関では、2022 年 11 月の ChatGPT のリリース以降、生成 AI を中心に AI への取組みが加速し、検討や導入に向けた動きが進展している。一方、金融機関においては、検討や利活用を進める過程で、リスク管理等の観点からの課題や留意点を認識している。

制度や規制などに関する動きも進展している。例えば、わが国の政府では、「AI 戦略会議」において議論が続けられており、また、本年 4 月に、経済産業省と総務省から「AI 事業者ガイドライン（第 1.0 版）」が公表されている。民間団体においても、金融機関の実務目線に立ったガイドライン等を作成する取組みが行われている。

（２） AI に関する FISC の活動

当センターでは、AI に関係する事項をテーマとするレポート類の作成・公表に加え、講演等の形で、情報発信を行っている。

最近の具体的な取組みとして、企業における AI の利活用について、倫理面での課題に関する議論が行われるようになっていることや、問題が指摘される事例が国内外で発生していることなどを踏まえ、AI と倫理に関する調査研究を行ってきたほか、生成 AI への関心の高まりを踏まえ、会員企業等との議論を重ねている。そして、そこから得られた情報や知見をもとに、生成 AI についての安全対策の視点からの調査研究や情報発信を行っている [1]。

---

[1] AI に関する、当センターの最近の取組事例は、次のとおりである。調査研究レポート及び News Letter は、会員向けに発信している。会員以外においては、当センターのホームページ経由で購入が可能である。

- 調査研究レポート「AI と倫理～AI 倫理原則・AI 倫理ガバナンスを通じた取組みの現状と課題」（2023 年 2 月）
- News Letter「AI と倫理～国際機関や主要国政府のガイドラインに関する動向調査～」（2022 年 6 月）
- 「金融機関による生成 AI の業務への利活用に関する暫定的考察」（2023 年 12 月）

（３） 本稿の位置づけと概要

本稿は、金融機関が AI を業務に利活用する動きが進展していることを踏まえ、代表的な利活用事例を紹介することと合わせ、AI の利活用において認識されている、及び今後認識することが必要となると考えられる課題と対応策に関する、当センターとしての考察を整理するものである。

昨年 12 月の「金融機関による生成 AI の業務への利活用に関する暫定的考察」（以下、「暫定的考察」という。）の公表以降における動向、具体的には、「AI 事業者ガイドライン」等の政府等における取組みの内容、また、当センターが引き続き実施している、会員企業等との議論から得られた情報等を踏まえ、暫定的考察の内容を改訂及び拡充している。その際、対象を、生成 AI を中心としつつ AI 全般に拡大しているほか、当センターが、金融情報システムに関して広く活用されているガイドラインである「金融機関等コンピュータシステムの安全対策基準・解説書」を策定していることを踏まえ、安全対策の観点に重点を置く形とし、表題を「金融機関による AI の業務への利活用に関する安全対策の観点からの考察」としている。

本稿の概要は、次のとおりである。まず、会員企業等との意見交換等から得られた情報をもとに、金融機関における AI の代表的な利活用事例を紹介する。次に、AI のうち生成 AI について、金融機関が認識する利活用上の課題と、それに対する当センターの考察を、暫定的考察の内容を再掲する形で整理する。そのうえで、金融機関における AI 全般の業務への利活用に当たっての課題と、それらに対して金融機関が取り組むべきと、当センターとして考える対応策を、改めて取りまとめる。最後に、AI に関する当センターの今後の活動予定を記載する。

なお、暫定的考察は、当センターの会員向けに情報発信したものであったところ、本稿は、会員に限定せず、金融情報システムの幅広い関係者向けに公表する。

１　金融機関における AI の代表的な利活用事例

当センターがこれまでに実施した調査、及び会員企業等との間で実施した意見交換等から得られた情報をもとに、AI の代表的な利活用事例を整理すると、以下のとおりである。大別すると、①ディープラーニング等を用いて、予測、提案又は決定を行うことができる「従来型 AI」の事例（事例１～３）と、②画像、文章等を生成する「生成 AI」の事例（事例４～８）である。

②のうち、事例４～６は、実際の利活用事例又は検討事例である一方、事例７及び８は、現時点で、将来的な利用可能性を見極めるための実証実験事例であり、実際の取組みとしては僅少である。なお、いずれも、事例をもとに表現を一般化したものである。

①「従来型 AI」の利活用事例

（事例１）　住宅ローン審査や法人顧客向けのオンライン完結型融資における活用。申込人の定量データや預金口座取引のデータを用いた分析・与信判断に AI を活用する。

（事例２）　消費性カードローン申込予測における活用。顧客属性や取引履歴、成約情報を学習さ

・ News Letter「米国金融機関等における生成 AI の利活用の動向」（2023 年 12 月）
・ News Letter「欧米金融機関における AI を活用したコールセンター業務の高度化」（2024 年 6 月）
・ 外部講演「金融機関における AI の利活用に向けた留意点～安全対策の視点から～」（2024 年 5 月）

せたAIモデルを用いて、顧客に最適な商品・サービスを提供するためのマーケティング予測データを抽出する。

（事例３） 保険金不正請求の早期検知における活用。契約の内容に加え、気象情報等の外部データを利用し、不正な請求を早期に検知する仕組みにおいてAIを活用する。

②「生成AI」の利活用事例

（事例４） 業者が提供する大規模言語モデル（LLM : Large Language Model）を利用したAIチャットサービスにより、業務を効率化する。具体的には、社内外向けの業務文書・メール文案の作成支援、文書作成時の言葉遣いや文法のチェック、海外とのコミュニケーションにおける各種言語への翻訳支援、チャットシステムとの対話を通じた企画・アイデア出し支援、分量が多い文献の要約や要点の整理、プログラムのコード生成支援、などである。

（事例５） プロンプトエンジニアリング（生成AIに対し、具体的で明確な指示や、必要なコンテキスト情報（付加情報）を与えることで、出力の精度や品質を高める技術）を用い、事業者が提供するLLMを使用したAIチャットサービスにより、社内外向けの業務文書やメール文案の作成業務を効率化する。具体的には、専用検索画面に質問のテンプレートを準備する、生成 AI への作業指示が組み込まれたボタンを用意する、などである。

（事例６） RAG（Retrieval-Augmented Generation（検索拡張生成）。生成AIに与える指示において、検索技術等を活用して得られた関連情報を追加し、必要な付加情報を与えることで、出力の精度や品質を高める技術）を、社内情報データベースとシステム連携して検索できる仕組み（システム）と組み合わせることで、業務文書の情報検索、照会対応業務、融資稟議書の作成業務を効率化する。

（事例７） LLM を、特定の業務やデータに適応させ、利用者の要件に合わせた出力が可能なモデルになるよう再学習させるプロセス（ファインチューニング）を経たうえで、営業支援、社内外向けの業務文書の作成等の業務に利用する。高い品質の特化型データや高度な技術的知識・スキルが求められるとされている。

（事例８） テキストだけではなく、音声・画像・動画など、異なる複数の種類の情報を組み合わせてAIで処理する技術を活用する。コールセンターでの電話応答の対応品質の向上など、様々な業務での活用が期待されている。

２ 生成AIの業務への利活用における課題と考察

当センターが昨年12月に公表した暫定的考察では、金融機関における生成AIの業務への利活用に関して、［図表］のとおり、課題を情報セキュリティ面、及び倫理面の別に分類したうえで、各々に対する考察を行っている。生成AIの利活用に当たっては、引き続きこうした点について留意が必要である。

［図表］生成 AI の利活用における課題及び考察

<table>
<tr><th colspan="3">課題</th><th>考察</th></tr>
<tr><td rowspan="2">情報セキュリティ面に関する課題</td><td>情報漏洩に関する課題</td><td>○　生成 AI にプロンプト入力された情報[2]が、生成 AI の機械学習において利用されることにより、他の情報と統計的に結びついた上で、また、正確又は不正確な内容で、生成 AI サービスから出力され、意図しない情報漏洩につながる可能性がある。</td><td rowspan="2">○　生成 AI を利用する場合、サービスの利用形態によっては、セキュリティ対策やデータ取扱いについて個別に設定することができないことがありうる。そうしたセキュリティ要件を満たすことが困難な場合においては、意図せざる情報漏洩を防止する観点から、生成 AI において機密性が高い情報を取り扱うことは適当ではない。
○　組織の許可を得ずに、職員等が生成 AI サービスを業務に利用する、いわゆる「シャドーIT」は、組織として誰がどのように使用しているか等の管理・統制がとれなくなる可能性があり、情報漏洩等のリスクが高まるため、適当ではない。
○　生成 AI の利用に当たっては、個別契約等に基づき、一定程度情報セキュリティ対策が施されたサービスを利用する場合であっても、当該サービスにおけるセキュリティ対策のほか、生成 AI の特性や固有リスク、取り扱う情報の内容を考慮することが必要である。
○　生成 AI 等の効率的で便利なツールの利用時には、インターネット翻訳サービス等と同様、その利便性がゆえに、職員は「機密性を有する情報を外部に送信している」という意識を</td></tr>
<tr><td>仕様・設定等に関する課題</td><td>○　生成 AI の利用に当たっては、事業者が提供するサービスの仕様を十分に理解したうえで利用することが必要となるほか、利用者が設定できる事項についても、その内容を十分に理解したうえで適切な設定等を行うことが必要となる。
事業者が提供する生成 AI サービスの場合、30 日監視保有[3]の仕組みや、クラウド環境が設けられるリージョン（海外／国内）などの各種仕様について十分な理解が求められる。そのうえで、オプトアウト[4]の設定や、プロンプトに対する制御方法（ルール化、フィルタリング、テンプレート等）[5]など、利用者側の設定等についても十分な理解にもとづく適切な設定等が求められる。また、外部サービス連携機能の利用や不特定多数の利用者への提供が想定されるサービスの性格上、セキュリティ対策やデータ取扱いについて個別対応ができない可能性もある点も、十分な確認が必要である。</td></tr>
</table>

[2] 生成 AI から回答を得るために、人が入力できる指示・命令文。
[3] 事業者が提供する LLM において、標準仕様として悪用監視のため、入出力データが 30 日間保有される仕組み。
[4] 事業者が提供する LLM においては、標準仕様として、入力データが機械学習される設定となっているところ、これを機械学習させない設定（オプトアウト機能）へ変更することが可能である。
[5] 質問をプロンプト入力する際に、質問方法のルールを決めたり、入力が禁止されている情報をフィルタリングしたりする仕組みが検討されているとされている。

倫理面に関する課題	モニタリングに関する課題	○　生成 AI にプロンプト入力された情報（指示・命令文）、学習に利用される情報（学習データ）、作成された情報（回答）のそれぞれにおける情報の正確性・正当性・妥当性、又は情報漏洩やプロンプトインジェクション [6]などについて、常時のモニタリングが望ましいが、人手による十分なモニタリングは困難である。	持ちづらく、機密情報を気軽に入力してしまいがちであると考えられている。こうした点を十分認識したうえで、サービスの利用形態や運用形態を検討することが必要である。
	虚偽情報が生成される課題	○　生成 AI において作成された情報に虚偽情報が含まれ、それが拡散し、社会等に影響を及ぼす可能性がある。特に「幻覚：ハルシネーション [7]」と呼ばれる根拠のない誤った内容が含まれる文章を作成してしまう可能性がある。	○　生成 AI による出力結果を用いる場合、当該結果をそのまま二次利用することが不適当（リスクがある）なケースがある。生成 AI の回答が正当か、ハルシネーションではないか等について、入念に確認する必要がある。 ○　生成 AI がどのような情報を用いて回答を作成したのかを、参照元を明示させるなど、人間が目で見て分かるようなフィードバックの仕組みを活用することは有効である。 ○　生成 AI の利活用に当たっては、利用する先の企業規模や生成 AI サービスの内容・影響度に応じて、AI 倫理ガバナンスに関する活動にリソースを適切に配分し、最適な AI 倫理ガバナンスの態勢整備を構築することが必要である。
	人権侵害に関する課題	○　生成 AI において作成された情報に個人情報や個人に関する差別や偏見、公平性等を侵害する情報 [8]が含まれ、それが拡散し個人の人権を侵害する可能性がある。	
	説明力・納得感に関する課題	○　生成 AI において作成された情報は、その情報（回答）を導くまでのプロセスを検証する仕組みが備わっておらず、納得性の高い説明や情報の正当性・妥当性等の判断が困難となる可能性がある。	

[6] 生成 AI に対して意図的に特殊な質問や命令を入力することで、意図しない結果を引き起こす、あるいは機密情報や公開すべきでないデータを取り出す手法。

[7] 生成 AI が事実に基づかない回答を生成する現象。その回答内容が、まるで幻覚：ハルシネーションのようにもっともらしい虚偽である特徴。

[8] プロンプト入力する際に、誤って個人情報を入力したため、生成 AI がその個人情報を学習することにより、当該個人情報に関するハルシネーションを生成し、虚偽性に起因する人権侵害が問題となるようなケースを想定。

３　金融機関における AI の業務への利活用に当たっての課題及び対応策

本稿におけるこれまでの記述を踏まえ、また、対象を、生成 AI を含む AI 全般に拡げたうえで、AI 提供者[9]及び AI 利用者[10]の双方の立場から、安全対策の観点に重点を置き、金融機関の課題と、それらに対して金融機関が取り組むべきと、当センターとして考える対応策を整理する。

具体的には、金融機関が AI を適切に利活用するに際しての課題を、（１）AI 利活用に係る方針策定及び態勢整備、（２）AI の適切な利用・運用管理、（３）AI に係る安全対策、及び（４）AI に係る教育、注意喚起等、の 4 つに分類し、それぞれについて、金融機関が取り組むべき対応策を記載している。なお、「２　生成 AI の業務への利活用における課題と考察」において示した、生成 AI に関する情報セキュリティ面、及び倫理面の課題、並びに取り組むべき対応は、以下に包含されている。

（１）AI 利活用に係る方針策定及び態勢整備

（金融機関の課題）

○　金融機関において、AI を試行的に利活用する段階を経て、利活用の業務の幅や利用者が拡がりつつある。こうしたなか、金融機関は、AI を利活用することで達成すべき目標や効果について、明確に認識することが必要である。そこでは、利活用する AI モデルや事業者が提供する AI サービスに関する十分な理解のほか、リスク管理や安全対策に関する適切な知識・判断が求められる。

○　そのうえで、金融機関は、AI を利活用する目的と、利活用に際しての方針を定め、それらを当該金融機関における組織としての共通認識とすることが必要である。

○　さらに、金融機関は、AI を適切に利活用するための態勢を整備し、利活用の方針と合わせ、AI 利活用に当たっての内外関係者の理解・信頼を得るため、説明責任を果たす責務がある。

（金融機関が取り組むべき対応策）

○　AI の利活用に係る方針を策定する。そこには、AI の利活用が社会全体に及ぼす影響を考慮し、リスク管理、安全対策等の観点を盛り込み、AI ガバナンスを実践するための行動規範や基本原則を定める。また、これらに適合した社内規程等を整備する。

○　組織として、役割分担、情報集約、情報提供を円滑かつ効果的に行うとともに、責任の所在を明確にし、適切なリスク管理、安全対策等を講じるために、AI の利活用に係る態勢を整備する。

○　AI の利活用に関して、内外関係者の理解・信頼・判断につながる情報の提供、説明を行う。

[9] 「AI 事業者ガイドライン」では、「AI 開発者が開発する AI システムに付加価値を加えて AI システム・サービスを AI 利用者に提供する役割を担う者」と定義している。

[10] 「AI 事業者ガイドライン」では、「AI 提供者から安全安心で信頼できる AI システム・サービスの提供を受け、AI 提供者が意図した範囲内で継続的に適正利用及び必要に応じて AI システムの運用を行う者」と定義している。

（２） AI の適切な利用・運用管理

（金融機関の課題）

○　金融機関は、AIに関する特性や前提を踏まえ、また、以下の点を認識したうえで、AIを適切に利用し、運用管理することが求められる。

- AIから生成・出力されたアウトプットには、①安全性、公平性、正確性、多様性・包摂性などの点において妥当性を欠く可能性、②著作権・知的財産権、業法免許・資格などに関する法令違反・権利侵害の内容を含む可能性、③正当性が確認できない、不正確な情報（ハルシネーション）となる可能性、などがあること。
- AIでは、データが重要な役割を果たしている。このため、データが正確でありかつ整合的であること、定期的に更新されていることなど、データの品質が確保されていること、加えて、データ内容や機密度等に応じたデータの管理方法が定められていることなど、データを適切に活用するための枠組みを整備する必要があること。
- AIモデル・AIサービスが急速な進化・進展を遂げつつある一方で、これらを巡る国内外の制度・法規制も変化している。AIモデル・AIサービスに関するその時点での正しい知識に基づき、制度・法規制に適合した利用・運用を行う必要があること。

○　金融機関における AI の利活用が適切であることについて、理解と信頼を得て評価されることが求められる。

（金融機関が取り組むべき対応策）

○　AI の適切な利用・運用管理方法を定める。AI で生成・出力されたアウトプットに関する留意点や確認、データの取扱い、システムの運用などにおいて、AI の特性等を理解したうえで、適切な利用・運用管理の方法を検討し、規定する。

○　AI の適切な利用・運用管理に当たっては、透明性が確保されたものとして規定する。

（３） AI に係る安全対策

（金融機関の課題）

○　AIの利活用に当たり、AIのセキュリティ確保が必要である。金融機関は、情報セキュリティに関するリスクを適切に管理し、顧客情報や取引情報等の安全性を確保するなど、自身が利用する情報システムに関して、安全性を確保することが求められる。

○　職員等が組織の許可を得ずに、適切な安全対策や利用状況の管理・統制がなされていない AI を業務に利活用することは、意図せず機密性・完全性・可用性を損なう、又は AI の安心安全な利活用が阻害される可能性があることから、情報漏洩等のリスクに対する適切な対処が求められる。

（金融機関が取り組むべき対応策）

○　利用者又は関係者に危害を及ぼすことがないように、AI に係る安全対策を講じる。

○　AI への入力及び生成・出力を通じて、情報漏洩等を生じさせないための、セキュリティに関する安全対策を講じる。

○　AI がサイバー攻撃、不正操作等を受け、意図しない振る舞いをすることにより、機密性・完全性・可用性が脅かされることがないよう、安全対策（AI モデル・AI サービスの安全対策の

確認を含む。）を講じる。 ○　AI サービス等の内容を踏まえ、AI サービス等の職員等の利用可否、利用形態を明確にする。

（４） AI に係る教育、注意喚起等

（金融機関の課題）
○　金融機関においては、AIに関する適切な利用・運用、安全対策を実現するために、利用者及び関係者が、AIに関する知識、リテラシー、倫理観等をもって、業務を遂行することが求められる。 ○　AI利活用の業務の幅や利用者が拡がりつつある点、及び技術・リスク管理等に関する各種情報が頻繁にアップデートされ得る点に留意したうえで、利用者及び関係者に、必要な情報提供及び教育を行い、認識を共有することが求められる。
（金融機関が取り組むべき対応策）
○　利用者及び関係者に、AI の正しい理解のもとで社会的に正しく利用する知識、リテラシー、倫理観等を持たせるために、教育及び注意喚起等を行う。

４　AI に関する当センターの今後の活動予定

当センターでは、AI に関して今後、以下の取組みを進める予定である。

（１） 会員企業等との意見交換等の継続

金融機関における AI の利活用の可能性、及び利活用に当たっての課題・取組みについて、会員企業等（金融機関、IT ベンダー等）との間の意見交換及び議論を継続し、情報収集を進めるとともに、知見を深める。

（２） 「金融機関等コンピュータシステムの安全対策基準・解説書」への AI 関連基準項目の追加

本稿での整理を踏まえ、「金融機関等コンピュータシステムの安全対策基準・解説書」において、AI を対象とする基準項目を追加する改訂を、今年度中に実施する。

（３） 生成 AI に関する利活用事例を中心とする調査研究レポートの公表

生成 AI に関する利活用事例の紹介を中心とする調査研究レポートを、今年度中に公表する。

以　上

【本件に関する照会先】
公益財団法人金融情報システムセンター（FISC）
企画部 fintech2@fisc.or.jp