

人工知能リスクマネジメントガイドラインに関する意見募集文書

目次

1. 序文	3
2. AI リスクマネジメントに対する MAS の監督アプローチ	4
3. ガイドラインの適用範囲	6
4. 提案される AIRG	7
5. 質問一覧	9
6. AI リスクマネジメントに関するガイドライン案	10

1. 序文

- 1.1. シンガポール金融管理局（MAS）は、金融機関（FI）における AI リスクのリスクマネジメントを強化し、金融セクターにおけるAIリスクマネジメントに関するMASの監督上の期待を定めるため、人工知能（AI）リスクマネジメントに関するガイドライン（以下「本ガイドライン」）¹の導入を提案している。本ガイドラインは、金融機関におけるAIリスクマネジメントの監督、主要なAIリスクマネジメントシステム・方針・手順、主要なAIライフサイクル管理、ならびにAI活用に必要な能力と体制に焦点を当てている。
- 1.2. 本ガイドラインは、金融セクター全体に一般的に適用可能な一連の期待事項を確立することを目的としており、規模やリスクプロファイルが異なる金融機関に対して比例的な方法で適用される可能性がある。本ガイドラインは、生成的AIを含む様々なAIアプリケーションや技術、ならびにAIエージェントなどの新たな開発にも一般的に適用されるべきである。しかしながら、MASはAIの進化する性質を認識しており、必要に応じて本ガイドラインを更新または補足する。
- 1.3. MASは、金融機関およびその他の関係者から本ガイドラインに関する意見を募集する。
- 1.4. 提出された意見は、提出者が明示的に公開を拒否しない限り、提出者名を明記した上で公表されることに留意されたい。したがって、提出者が以下のいずれかを希望する場合：
 - (a) 提出内容の全部または一部（提出者の身元を除く）、または
 - (b) 提出内容全体と提出者の身元をMASへの提出書類に明示的に記載すること。MASは匿名でない提出書類のみを公開する。さらに、MASは、提出書類が誹謗中傷的または攻撃的であると判断される場合など、公開が公共の利益に反すると考える場合、提出書類を公開しない権利を留保する。
- 1.5. 意見募集文書への書面によるコメントは、**2026年1月31日**までに以下のリンクより提出すること：
<https://form.gov.sg/690b2a3b024ee5eebbfcf7f1>

¹ 本諮問文書におけるAIの範囲には、機械学習、深層学習、強化学習技術に基づくAI、ならびに生成的AI、AIエージェント、およびガイドライン第1.2項で定めるAIの提案範囲に該当する新たなAI技術が含まれる。

2. AI リスクマネジメントに対する MAS の監督アプローチ

- 2.1. 金融分野における AI の利用は新しいものではない。しかし、最近の AI 技術の進歩により、金融分野での AI 活用への関心が高まっている。AI は事業や機能分野全体に大きな利益をもたらす可能性がある一方で、特に生成的 AI や AI エージェントといった新しく複雑な AI 技術の利用は新たな課題をもたらす。金融機関（FI）内の事業領域や機能領域で AI 導入が普及するにつれ、既存リスクが増幅したり新たなリスクが生じたりする可能性がある。例えば、リスクアセスメントに用いる AI モデルの性能不良は多大な財務損失を招き、AI システムの予期せぬ動作は重要業務を混乱させ、顧客対応 AI システムの不適切な出力は顧客への損害や金銭的損失をもたらす恐れがある。
- 2.2. 生成的 AI は、理解が不十分で緩和が困難なリスクも生み出す。例えば、説得力はあるが虚偽の情報を生成する幻覚現象、より複雑な手法の使用から生じる予測不能な動作、意思決定プロセスの説明における根本的な課題などである。その他のリスクには、プロンプト・インジェクション攻撃などのセキュリティ脆弱性、サードパーティサービス利用時のデータ漏洩リスク、著作権保護コンテンツでの学習や既存著作権侵害の可能性のある出力生成による知的財産権侵害、少数の主要生成的 AI プロバイダへの過度の依存による集中リスク、生成 AI への過度の依存から生じる人的要因リスクなどが含まれる。
- 2.3. 生成的 AI を活用する AI エージェントのようなさらに新しい技術の使用は、より大きな自律性とツールへのアクセス能力を伴うため、さらに重大なリスクをもたらす可能性がある。金融機関の内部システムへのアクセス権限を与えられた AI エージェントは、事業目標や顧客の利益と整合しない行動を自律的に実行する可能性がある。また、侵害された AI エージェントは機密データを流出させたり、悪意のあるコマンドを実行したりする恐れがある。
- 2.4. 金融セクターにおける AI の利用拡大と関連リスクの高まりを受け、MAS は金融機関が AI を責任を持って利用するための指針となる基本原則を確立した。2018 年には、金融業界と共同で「公平性、倫理、説明責任、透明性（FEAT）」原則⁽²⁾を策定し、責任ある AI・データ分析の展開を推進した。金融機関による FEAT 実施を支援するため、MAS は 2019 年 11 月、業界コンソーシアムと共同で「ベリタス・イニシアチブ」³を開始した。同イニシアチブは、金融機関が AI・データ分析ソリューションに FEAT 原則を組み込むことを支援する目的で、アセスメント手法、ツールキット、関連事例研究を公開している。
- 2.5. 生成的 AI の台頭を受け、そのリスクと機会を検討するため、プロジェクト・マインドフォージ⁽⁴⁾が設立された。プロジェクト・マインドフォージの第 1 フェーズは銀行コンソーシアムが主導し、2023 年 11 月に生成的 AI 向けリスク枠組みを発表した。第 2 フェーズではコンソーシアムを拡大し、資本市場や保険など金融セクターの他分野の金融機関も参加させた。拡大コンソーシアムは業界主導の「AI リスクマネジメントハンドブック」の作成を進めており、これは金融機関がガイドライン⁵を実施する際の補助ガイドとなる。
- 2.6. MAS はまた最近、金融機関による AI 利用の観点、および金融機関に対するサードパーティによる AI 利用の観点から、AI および生成的 AI に関連する情報文書を発表した：
 - (a) 生成的人工知能に関連するサイバーリスク（2024 年 7 月）⁶。本資料では、生成的人工知能から生じる主要なサイバー脅威の概要、リスクへの影響、および金融機関がこうしたリスクに対処す

² <https://www.mas.gov.sg/publications/monographs-or-information-paper/2018/feat>

³ <https://www.mas.gov.sg/schemes-and-initiatives/veritas>

⁴ <https://www.mas.gov.sg/schemes-and-initiatives/project-mindforge>

⁵ プロジェクト「Mindforge AI リスクマネジメントハンドブック」は 2026 年 1 月までに公開される。

⁶ <https://www.mas.gov.sg/regulation/circulars/cyber-risks-associated-with-generative-artificial-intelligence>

るために講じ得る緩和策について概説した。また、ディープフェイク、フィッシング、マルウェアなど生成的人工知能によって可能となる領域や、データ漏洩やモデル操作など展開済み生成的人工知能に対する脅威についても扱った。

- (b) 銀行における AI モデルリスク管理 (MRM) (2024 年 12 月)⁷。本論文は、AI の堅牢な監視とガバナンス、包括的な AI 識別・リスク重要性評価・インベントリのための主要なリスクマネジメントシステムとプロセス、AI の厳格な開発・妥当性確認・展開のための基準と統制に関する、銀行における優れた AI MRM 実践例を強調した。
- (c) ディープフェイクに関連するサイバーリスク (2025 年 9 月)⁸。本論文は、ディープフェイクがもたらす新たな脅威とリスクの概要、金融セクターへの潜在的影響、および金融機関がこうしたリスクに対処するために実施可能な緩和策について述べた。本論文は、ディープフェイクが金融セクターに影響を与える 3 つの主要領域を扱った。これには、生体認証の回避、なりすましや詐欺のためのソーシャルエンジニアリングの実施、誤情報・偽情報の拡散促進におけるディープフェイクの利用が含まれる。

- 2.7. 金融セクターの取り組みは、国家レベルでの主要なイニシアチブによっても支援されている。例えば、情報通信メディア開発庁 (IMDA) のモデル AI ガバナンス枠組み⁹ や、AI 検証財団 (AI Verify Foundation) の各種イニシアチブ¹⁰ などである。
- 2.8. 本ガイドラインは、FEAT 原則、業界との協働経験、および金融庁 (MAS) の AI リスク関連情報文書を基盤としている。FEAT が AI 利用における公平性、倫理、説明責任、透明性の原則を定めるのに対し、本ガイドラインは金融セクターにおける AI 利用時のリスクマネジメントに関する高水準の監督上の期待を明確化することに焦点を当てる。本ガイドラインは、金融機関における AI リスクマネジメントの監督、主要な AI リスクマネジメントシステム・方針・手順、主要な AI ライフサイクル管理、AI 活用に必要な能力・体制に焦点を当てる。¹¹
- 2.9. 本ガイドラインは、生成的 AI を含む様々な AI アプリケーションや技術、ならびに AI エージェントなどの新たな開発にも一般的に適用されるべきである。しかしながら、MAS は AI の進化する性質を認識しており、必要に応じて本ガイドラインを更新または補足する。

⁷ <https://www.mas.gov.sg/publications/monographs-or-information-paper/2024/artificial-intelligence-model-riskmanagement>

⁸ <https://www.mas.gov.sg/regulation/circulars/cyber-risks-associated-with-deepfakes>

⁹ <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf>

¹⁰ <https://aiverifyfoundation.sg/>

¹¹ 外部主体による AI の利用から生じる可能性のある金融機関へのその他のリスク、例えば AI を活用したサイバー攻撃や詐欺などがある。これらは本ガイドラインの範囲外であり、2.6(a)項および 2.6(c)項で言及されているものなど、MAS の他の刊行物で扱われている。

3. ガイドラインの適用範囲

- 3.1. MAS は、本ガイドラインを全ての金融機関（FI）に適用することを提案する。¹² シンガポールに所在する金融機関で、他国に親事業体を有する支店または子会社である場合は、当該親事業体の AI リスクマネジメント枠組みが本ガイドラインで定められた期待を満たす限り、その枠組みを活用することができる。
- 3.2. ¹³MAS は、AI が多様なユースケースに適用可能であり、そのリスクは金融機関の規模、範囲、ビジネスモデルによって異なることを認識している。金融機関は、自らの活動規模・性質、AI の利用状況、リスクプロファイル、および本ガイドラインが特定の AI モデル・システム・ユースケースに与える関連性に応じて、比例的な方法で本ガイドラインを実施することができる。
- 3.3. 本ガイドラインの目的上、AI の提案範囲には、以下のように定義される AI モデル、システム、またはユースケースが含まれる：
- (a) モデルとは、仮定や入力データを推定値、決定、推奨事項などの出力に変換する手法またはアプローチを指す。
 - (b) システムは、一つ以上のモデルやその他の機械ベースの構成要素で構成されることがある。
 - (c) ユースケースとは、モデルやシステムが適用される特定の現実世界の状況を指す。
 - (d) AI には、入力から学習および／または推論を行い、物理的または仮想的環境に影響を与える可能性のある推定値、予測、コンテンツ、要約、推奨事項、決定などの出力を生成するモデルやシステムを含むユースケースが含まれる。これらは展開後の自律性や適応性のレベルが異なる。¹⁴ 出力が事前に定義されたプログラミングロジックやルールのみに基づく計算ツールやツールは、本ガイドラインの目的上、AI とはみなされない。

質問 1. MAS は、ガイドラインのすべての金融機関への比例的な適用、およびガイドライン第 1.5 項と附属書に定める比例的適用に関するガイダンスについて意見を求めます。

質問 2. MAS は、本ガイドラインの適用対象となる AI ユースケース、システム及びモデルの提案範囲について意見を求めます。

¹² 2022 年金融サービス・市場法第 2 条に定義される通り。

¹³ 全ての金融機関は、AI 導入レベルに応じた基本方針を策定すべきである。これらの方針は、AI 利用の監督責任者、AI の許容・禁止用途に関するガイドライン、および当該ガイドラインの伝達・検証・見直しを規定するものとする。AI 監督および主要な AI リスクマネジメント・機構・方針・手順に関する期待事項は、AI を業務プロセスに統合して利用する金融機関にのみ適用される。AI ライフサイクル管理、および AI 利用のための能力・キャパシティに関する期待事項は、金融機関における AI ユースケース、システム、モデルの関連性およびリスクの重要性に基づいて適用される可能性がある。これらのガイドラインを比例原則に基づいて実施する詳細については、ガイドラインの 1.5 項および附属書に規定されている。

¹⁴ 本ガイドラインの目的上、これには一般的に機械学習、深層学習、強化学習技術に基づく AI、生成的 AI、AI エージェント、およびその他の新たな AI 技術が含まれる。

4. 提案される AIRG

AI 監督

- 4.1. 本ガイドラインは、金融機関の取締役会及び上級管理職が、様々な分野における AI 関連リスクを統治・監督することについて、MAS の期待を定めている。これには、金融機関における AI の使用を識別・把握し、そのリスクの重要性を評価し、必要なガバナンス及びリスクマネジメントの枠組み、方針、プロセスを整備し、AI のライフサイクル全体を通じてリスクを管理するための枠組み、構造、方針及びプロセスの確立と堅固な実施が含まれる。取締役会及び上級管理職は、AI 利用に関する適切なリスク文化を醸成するとともに、既存の組織全体のリスクマネジメントが AI 利用によって生じるリスクに対応できるよう更新されることを確保すべきである。金融機関の AI リスクエクスポージャー全体が重要とみなされる場合¹⁵、MAS は、当該金融機関が適切な監督を確保し、リスクマネジメントにおける潜在的なギャップに積極的に対処するため、専任の部門横断的委員会を設置することを提案する。

質問 3. MAS は、AI リスクマネジメントの監督における取締役会及び上級管理職の提案された責任について意見を求めます。

質問 4. 金融機関の AI リスクエクスポージャーが全体として重要とみなされる場合、AI リスクを監督する専門の部門横断委員会を設置する提案について、また組織レベルでそのような AI リスクエクスポージャーをどのように評価すべきかについて、意見を求めます。

主要な AI リスクマネジメントシステム、方針及び手順

AI の特定

- 4.2. 金融機関内で AI が使用されている箇所を識別することは、そのような AI の使用に対して適切なガバナンス、リスクマネジメント標準、および統制を効果的に適用するための重要な前提条件である。したがって、MAS は金融機関が、この識別プロセスを促進し、関連するすべての事業および機能領域にわたって AI の使用を一貫して特定することを確実にするために、堅牢なシステムによって支えられた明確な定義、規準、およびプロセスを確立することを期待している。

AI インベントリ

- 4.3. 承認されていない AI の使用、特にリスクの高いユースケースでは、意図しない結果を招き、金融機関がリスク許容度を超える AI リスクに晒される可能性がある。このリスクを緩和するため、MAS は金融機関に対し、ガバナンスと監視、ならびに AI ライフサイクル全体を通じたリスクマネジメントを支援するため、AI ユースケース、システム、モデルに関する正確かつ最新のインベントリを確立・維持することを提案する。このインベントリは、AI 専用に構築するか、既存のインベントリを強化して構築することができる。いずれの場合も、AI インベントリと金融機関内の他の関連インベントリとの間に明確な連携が存在するべきである。

¹⁵ 金融機関は、ガイドライン第 3.8 項から第 3.11 項に定めるリスク重要性評価に関するガイダンスに基づき、自社の AI ユースケースの重要性と AI リスクへの全体的なエクスポージャーを評価すべきである。金融機関の AI に対する全体的なリスクエクスポージャーの重要性は、ガイドライン第 3.10 項に定める次元に加え、金融機関の事業戦略及び全体的なリスクプロファイルへの影響を考慮すべきである。一般的に、金融機関において、重要な事業部門や機能領域への高リスク AI の展開などにより、金融機関またはその顧客に悪影響を及ぼす可能性のある重大なリスクに晒される AI ユースケースが一つ以上存在する場合は、当該金融機関は、管理監督強化（すなわち、専任の部門横断的委員会の設置）が必要かどうかを評価すべきである。

リスク重要性アセスメント

- 4.4. MAS は、AI が多様なリスクレベルを伴う幅広い業務・機能領域で使用されていることを認識している。MAS は、金融機関が AI のユースケース、システム、モデルにおけるリスクの重要性を評価するための適切なアセスメント手法を導入することを期待する。MAS は、リスク重要性アセスメントが少なくとも影響度、複雑性、依存度の主要な側面をカバーすることを提案する。

質問 5. MAS は、金融機関が、堅牢なシステムに支えられた明確な定義、規準、プロセスを確立し、関連するすべての業務・機能領域における AI の使用を一貫して特定することを容易にするという提案について、意見を求めます。

質問 6. 金融機関が、全ての AI 利用状況について正確かつ最新のインベントリを確立・維持する提案について、MAS は意見を求めます。

質問 7. 金融機関が AI リスク重要性アセスメントにおいて把握すべき影響度、複雑性、依存度という提案されたリスク次元について、また他に含めるべきリスク次元があるかについて、MAS は意見を求めます。

AI ライフサイクル管理

- 4.5. MAS は、金融機関が AI モデル、システム、またはユースケースとの関連性に基づき、かつ特定の AI モデル、システム、またはユースケースの評価されたリスク重要性に比例した、AI ライフサイクル全体をカバーする堅牢な管理策を計画・実施することを期待している。金融機関が AI モデル、システム、またはユースケースとの関連性を評価し、比例的な方法で適用すべき主要領域には、データ管理、公平性、透明性と説明可能性、人的監視、サードパーティ AI リスクの管理、AI の選定、評価とテスト、技術とサイバーセキュリティ、再現性と監査可能性、レビュー、モニタリング、変更管理が含まれる。

質問 8. MAS は、AI ライフサイクル全体に適用すべき提案された標準、プロセス、統制、および金融機関が AI モデル、システム、ユースケースとの関連性を評価し、比例的な方法で適用すべき主要分野について意見を募集している。

一般

- 4.6. 上記で強調した分野に加え、MAS はガイドラインのその他の側面に関する意見も歓迎する。

質問 9. MAS は、これまでの質問で取り上げられていないガイドラインのあらゆる側面、および提案されたガイドラインでカバーされていない AI リスクマネジメントの側面について、意見を募集する。

実施

- 4.7. MAS は、金融機関（FI）によって AI リスクマネジメントの実践の成熟度が異なることを認識している。したがって、MAS は、ガイドライン発行後 12 ヶ月の移行期間を設けることを提案する。これにより、FI はガイドラインを適切に評価し、実施することができる。

質問 10. MAS は、提案されている 12 ヶ月の移行期間について意見を募集する。

5. 質問一覧

S/N	質問	ページ
質問 1	MAS は、ガイドラインのすべての金融機関への比例的な適用、およびガイドラインの 1.5 項と附属書に記載された比例的適用に関するガイダンスについて、意見を募集している。	7
質問 2	MAS は、ガイドライン適用における AI ユースケース、システム及びモデルの提案範囲について意見を求めます。	7
質問 3	MAS は、AI リスクマネジメントの監督における取締役会及び上級管理職の提案された責任について意見を募集する。	7
質問 4	MAS は、金融機関の AI リスクエクスポージャーが全体として重要とみなされる場合、金融機関が AI リスクを監督するための専任の部門横断委員会を設置する提案について意見を求めるとともに、組織レベルでそのような AI リスクエクスポージャー全体をどのように評価すべきかについて意見を求めます。	8
質問 5	MAS は、金融機関が、関連する全ての事業分野及び機能分野において AI を一貫して特定することを容易にするため、堅牢なシステムに裏打ちされた明確な定義、規準及びプロセスを確立する提案について意見を求めます。	8
質問 6	MAS は、金融機関が AI の使用状況に関する正確かつ最新のインベントリを確立・維持する提案について意見を募集している。	8
質問 7	MAS は、金融機関が AI リスクの重要性アセスメントにおいて把握すべき影響度、複雑性、依存度という提案されたリスク次元について、また他に含めるべきリスク次元があるかどうかについて、意見を募集している。	8
質問 8	MAS は、AI ライフサイクル全体に適用すべき標準、プロセス、統制に関する提案について意見を求めるとともに、金融機関が AI モデル、システム、ユースケースとの関連性を評価し、比例的な方法で適用すべき主要領域について意見を求めます。	9
質問 9	MAS は、これまでの質問で取り上げられていないガイドラインのあらゆる側面、および提案されたガイドラインで取り上げられていない AI リスクマネジメントの側面について、意見を募集している。	9
質問 10	MAS は、提案されている 12 ヶ月の移行期間について意見を求めます。	9

6. AI リスクマネジメントに関するガイドライン案

1 はじめに

1.1 人工知能（AI）リスクマネジメントに関するガイドライン（AIRG）は、金融機関（FI）における AI リスクマネジメントに関する MAS の監督上の期待を定めたものである。¹⁶ 金融セクターにおける AI の利用を指導する既存の公平性、倫理、説明責任、透明性（FEAT）原則¹⁷ は引き続き適用される¹⁸。本ガイドラインは、監督、主要な AI リスクマネジメントシステム、方針、手順、AI ライフサイクル管理、AI 利用のための能力とキャパシティの分野における MAS の監督上の期待を定めることで、FEAT 原則を補完するものである。本ガイドラインで扱う主要セクションの概要は下記の通りである。

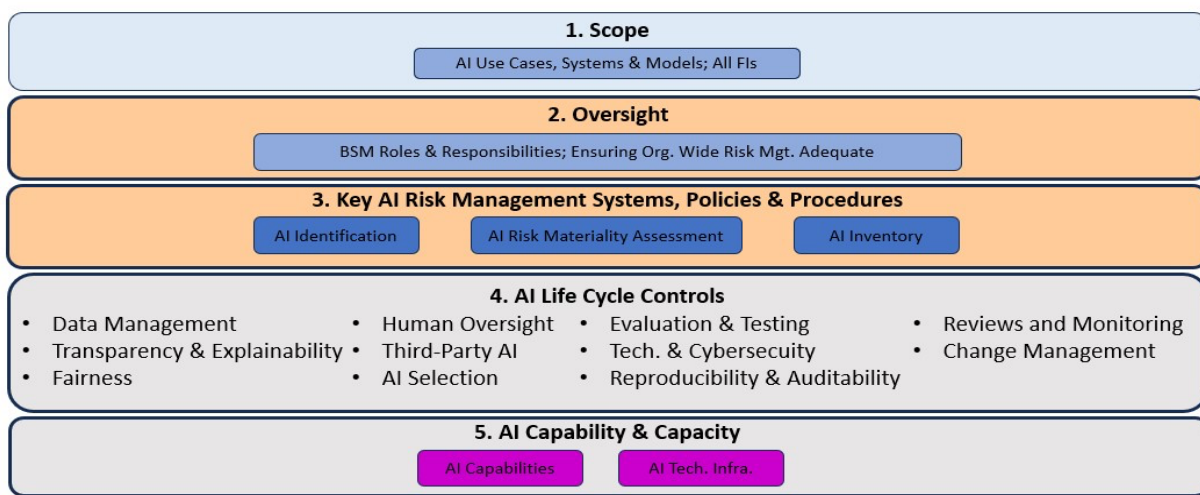


図 1 : AIRG の概要

- 1.2 本ガイドラインにおいて、AI とは AI モデル、システム、またはユースケースを指す場合がある。¹⁹ これらは以下のように定義される：
- モデルとは、仮定や入力データを推定値、決定、推奨事項などの出力に変換する手法またはアプローチを指す。
 - システムは、1 つ以上のモデルとその他の機械ベースの構成要素で構成されることがある。
 - ユースケースとは、モデルやシステムが適用される具体的な実世界の状況を指す。
 - AI には、入力から学習および／または推論を行い、物理的または仮想的環境に影響を与える可能性のある推定値、予測、コンテンツ、要約、推奨事項、決定などの出力を生成するモデルやシステムを含むユースケースが含まれる。これらは展開後の自律性や適応性のレベルが異なる（²⁰）。

¹⁶ 2022 年金融サービス・市場法第 2 条に定義される通り。

¹⁷ <https://www.mas.gov.sg/publications/monographs-or-information-paper/2018/feat>

¹⁸ 金融機関は、AI 活用の指針として、国家レベルの主要な取り組み（例：情報通信メディア開発庁のモデル AI ガバナンス枠組み）や、AI 検証財団（AI Verify Foundation）の各種イニシアチブも参照できる。

¹⁹ 特定の AI ユースケース、AI システム、AI モデルに関連する事項については、本ガイドラインにおいて該当する用語を明示的に使用する。

²⁰ 本ガイドラインの目的上、これには一般的に機械学習、深層学習、強化学習技術に基づく AI、生成的 AI、AI エージェント、およびその他の新たな AI 技術が含まれる。

事前定義されたプログラミングロジックやルールのみに基づいて出力を生成する計算機やツールは、本ガイドラインの目的上、AIとは見なされない。

- 1.3 本ガイドラインは、生成的 AI や AI エージェントなどの新技術を含む AI 技術を展開する際に、全ての金融機関が遵守すべき高水準の期待事項を定める。シンガポールに所在する支店・子会社で親事業体が他国にある金融機関は、親事業体の AI リスクマネジメント枠組みを活用して本ガイドラインの期待事項を満たすことが可能である。
- 1.4 金融機関は、自らの事業規模・性質、および AI 利用がもたらす重大なリスクの可能性に応じて、本ガイドラインを実施すべきである。AI は多様な用途に適用可能であり、そのリスクは金融機関の規模・範囲・ビジネスモデルによって異なる。
- 1.5 全ての金融機関は、最低限、自社の AI 導入レベルに見合った AI 利用の基本方針を策定すべきである。この基本方針では、AI 利用の監督責任者、AI の許容・禁止利用に関するガイドライン、および当該ガイドラインの伝達・検証・見直しについて定める必要がある。AI を業務プロセスに統合して利用する金融機関は、最低限、AI 利用を監督する枠組み・方針・手順を確立すること、AI のユースケース・システム・モデルについて明確な識別と堅牢なリスク重要性アセスメントを実施すること、適切な AI インベントリを整備することが求められる。こうした監督体制や AI リスクマネジメントシステム・方針・手順の程度は、金融機関の業務規模・性質、AI の利用状況、リスクプロファイルに応じて比例的に設定されるべきである。業務プロセスに統合された AI を利用する金融機関に対する期待事項は、本ガイドラインの第 2 節及び第 3 節に規定されている。金融機関における業務プロセスに統合された AI の利用を構成する要素、及び基本方針の範囲に関する詳細と事例は、附属書に示されている。
- 1.6 AI ライフサイクル標準と管理策の適用、ならびに AI 利用のための能力構築と技術インフラの整備（本ガイドライン第 4 節 および第 5 節に規定）は、金融機関における AI ユースケース、システム、モデルとの関連性に基づき調整される。これらの要素が AI ユースケース、システム、モデルに関連する場合、金融機関はリスクの重要性に基づいて実施する。例えば、MAS は、顧客またはリスクマネジメントの結果に重大な影響を与える方法で AI を利用する金融機関⁽²¹⁾ に対し、より強固な AI ライフサイクル標準と管理、ならびに高度な能力と技術インフラを期待する。透明性、説明可能性、公平性に関連する AI ライフサイクル標準や管理は、信用スコアリング、保険引受、金融助言の提供、ファンド管理サービスなど顧客に影響を与える分野で利用される AI にとってより関連性が高い場合もある。意思決定ではなく人間を支援する目的で低リスク重要性領域で使用される AI（例：文書作成支援のコパイロット）については、AI アルゴリズムの選定やストレステストなどの特定評価・試験手法に関連する AI ライフサイクル標準・統制は重要度が低い可能性がある。ただし、データ管理、安全性、サイバーセキュリティに関連する AI ライフサイクル標準・統制は重要であり、比例原則に基づき適用すべきである。
- 1.7 AI の発展のスピードを考慮し、MAS は金融機関に対し、AI の発展を踏まえた AI リスクマネジメントの取り組みの適切性を定期的に見直し、こうした発展によって生じる可能性のある新たな AI リスクや顕在化した AI リスクに対処することを期待している。

²¹ これには、規制資本の計算、規制報告、主要な財務リスク・業務リスク・マネーロンダリング防止リスク・不正リスクの管理、あるいは顧客成果に重大な影響を及ぼす規制対象業務（例：金融アドバイザーサービスの提供やファンド運用）の実施などが含まれる可能性がある。MAS は、影響の程度がこれらの分野における AI の活用方法に依存することを認識しており、金融機関が評価されたリスクの重要性に基づいて比例した標準と管理を適用することを期待している。

リスク

- 1.8 AIの利用は事業分野や機能分野におけるパフォーマンス向上に寄与し得るが、その複雑性と確率的性質は不確実性の増大、ならびに単純な手法の利用と比較して識別が困難な予期せぬ行動やより偏った行動を引き起こし得る²²。AIの複雑性の増大は、不正確またはバイアスのかかった出力結果の理解と説明における課題も招く。したがって、AI利用から生じ得る一般的なリスクには以下が含まれる：
- a. 財務リスク：例えばリスクマネジメントにAIを使用する場合、AIの不確実性の増大や予期せぬ動作が、不適切なリスクアセスメントやそれに伴う財務損失につながる可能性がある。
 - b. 業務リスク：金融機関の業務自動化にAIを使用した場合、その予期せぬ動作が業務中断や重要プロセスにおけるエラーを引き起こす可能性がある。
 - c. 行動リスク：例えば、AIの使用が偏った結果を生み、特定の顧客グループへの不公平な扱い、特定の投資タイプへのバイアス、金融機関と顧客間の利益相反を引き起こす可能性がある。
 - d. 金融犯罪リスク：例えば、マネーロンダリング対策を支援するためにAIを使用する場合、AIの不確実性や予期せぬ動作が大きければ、疑わしい取引が検出されない可能性がある。
 - e. 評判リスク。例えば、チャットボットなどの顧客対応システムが誤った情報を提供したり、不快な発言をしたりすることで、メディアの否定的な注目や評判の低下を招く可能性がある。
- 1.9 生成的 AI では、AIに関連する既存のリスクが増幅される可能性がある。生成的 AI の複雑性が高いため、従来のAIと比較して不確実性と予期せぬ行動がさらに増大する。生成的 AI の入力・出力の非構造化特性と、この分野における確立された手法の不足も、生成的 AI の行動や出力を評価・テストし、理解・説明することを困難にしている。生成的 AI のトレーニングに用いられる多様でしばしば不透明なデータソースと、生成的 AI 出力のバイアス評価の困難さが相まって、顧客に不公平な結果をもたらす決定につながる可能性もある。生成的 AI はまた、以下のような相互に関連する様々なリスクを引き起こす可能性がある：
- a. セキュリティリスク：プロンプト・インジェクションやデータ・ポイズニングによる生成的 AI システムへの敵対的攻撃など。²³
 - b. プライバシーリスク：例えば、サードパーティの生成的 AI 製品・サービスの利用による機密情報・顧客データの漏洩、または適切な同意なしに生成的 AI で顧客データが不適切に使用されることによるリスク。
 - c. 法的・知的財産リスク：既存の著作権や特許を侵害するデータで訓練された生成的 AI の使用に起因するリスク。
 - d. サードパーティリスク。例えば、ミッションクリティカルな領域で少数の支配的な生成的 AI プロバイダへの過度の依存、あるいはセキュリティ管理が不十分なオープンソースモデルの使用から生じるリスク。
 - e. 運用リスク。例えば、生成的 AI の停止によるサービス中断の可能性や、展開前のテストが不十分であることに起因する問題など。

²² こうした不確実性や行動は、訓練データに内在するノイズ、代表者に欠ける訓練データ、あるいはモデルが訓練データに存在しないシナリオに遭遇した場合に生じうる。また、時間の経過に伴うデータやモデルの変化によっても生じうる。

²³ こうしたリスクの詳細については、MASの「生成的人工知能に関連するサイバーリスクに関する情報文書」で広く扱われている。同文書は <https://www.mas.gov.sg/regulation/circulars/cyber-risks-associated-with-generative-artificial-intelligence> で閲覧可能だ。

- f. 人的要因リスク。例えば、ミッションクリティカル領域で使用される生成的 AI に対する人間の監督が不十分であること、または生成的 AI への過度の依存によるスキルの低下から生じるリスク。
- 1.10 同様に、より大きな自律性とツールへのアクセス権限を与えられる可能性のある AI エージェントなどの新技術の利用は、これらのリスクをさらに増幅させる恐れがある。例えば、ツールへのアクセス権を持つ AI エージェントは、人間の目標と AI エージェントがそれを行動に変換する方法との間に乖離があるため、金融機関の事業目標や顧客の最善の利益に沿わない行動を自律的に実行する可能性がある。内部システムや外部ツールへのアクセス権を持つ侵害された AI エージェントは、機密データの流出や大規模な悪意のあるコマンドの実行に利用され、セキュリティリスクを増幅させる恐れがある。
- 1.11 MAS は AI アプリケーションの継続的な成長を認識しており、必要に応じて本ガイドラインを更新または補強する。

2 AI の監視

- 2.1 金融機関の取締役会（以下「取締役会」）及び上級管理職は、金融機関全体における AI リスクマネジメントを支援する堅牢な枠組み、方針及び手順の確立と監督において重要な役割を担う。
- 2.2 **取締役会及び上級管理職は、AI 関連リスクの効果的な監督を維持し、AI 利用に適したリスク文化を醸成するとともに、AI の利用が他の監督上の期待を満たす能力と矛盾しないことを確保すべきである²⁴。**これには、以下の目的のための強固な枠組み、構造、方針及びプロセスの確立と実施が含まれる：
- a. AI のユースケース、システムまたはモデル（自社開発およびサードパーティの AI を含む）を識別すること²⁵。
 - b. AI 関連リスクの重要性を評価すること；
 - c. AI のユースケース、システム、モデルの一覧を管理し、事前に定義されたリスク許容度に基づいてその使用を統制すること；
 - d. AI ユースケース、システム、モデルをライフサイクル全体を通じて管理する。
 - e. 金融機関における AI ユースケース、システム、モデルの開発・展開に必要な能力と体制を構築する。
- 2.3 ²⁶**取締役会及び上級管理職は、組織全体の既存のリスクマネジメント枠組み、方針、慣行が、AI がもたらすリスクを適切に識別、評価、対処することを確保すべきである。**組織全体の既存のリスクマネジメント領域において、金融機関は以下を行うべきである：
- a. 関連する全ての AI リスクを識別し、アセスメントすること；
 - b. 当該リスクに対処するため、関連する方針と手順を更新すること；
 - c. 当該リスクの緩和のための適切な戦略と統制を導入すること；

²⁴ 例えば、金融機関は AI を活用して製品やサービスを提供する場合でも、MAS の「公正取引に関するガイドライン」を遵守し続けるべきだ。

²⁵ 本ガイドラインにおいて、サードパーティ AI とは、サードパーティシステム、モデル、AI に使用されるデータを含む、サードパーティ AI 製品・サービスの全てのプロバイダを指す。AI が導入された既存のサードパーティ製品・サービスもこれに含まれる。

²⁶ AI リスクが関連する主要な既存のリスクマネジメント領域には、モデルリスク、オペレーショナルリスク、レピュテーションリスク、データリスク、テクノロジー・サイバーセキュリティリスク、サードパーティリスク、法務・コンプライアンスリスク、財務リスク、行動リスク、環境リスクなどが含まれるが、これらに限定されない。

- d. 当該リスクに関する金融機関のリスク許容度を定義すること；
- e. 当該リスクに対する関連指標と適切なリスク許容度閾値を設定すること；
- f. 当該指標及びリスク許容度閾値の遵守状況を監視する；
- g. 異なる事業部門や機能間における AI リスクマネジメントの役割と責任を明確に定めること；
- h. 当該リスク許容度閾値の超過及び AI 関連インシデントについて、取締役会及び上級管理職への報告に関する明確な方針と手順を定めること；並びに
- i. 新たな AI 技術の発展、金融機関のリスクプロファイルや事業戦略の変化、AI 規制の動向を反映させるため、定期的な見直しを実施すること。

2.4 **取締役会及び上級管理職は、AI リスクマネジメントにおいて金融機関全体で一貫した標準、明確な説明責任、強固な連携を確保すべきである。**金融機関は、全ての AI 関連リスクを管理する新たな集中型機能の設置や、AI に起因する漸増リスクを既存のリスク管理機能に割り当て管理する分散型アプローチなど、様々な AI リスクマネジメント手法を採用し得るが、取締役会及び上級管理職は依然として AI リスクの一貫した連携管理を確保すべきである。金融機関の AI リスクへの全体的なエクスポージャーが重要と判断される場合²⁷、金融機関は適切な監督を確保し、リスクマネジメントのカバー範囲における潜在的なギャップに積極的に対処するため、専用の部門横断型委員会を設置すべきである。取締役会と上級管理職はまた、AI が導入された場合でも、金融機関がこれらのリスク領域に関連する既存の規制要件をすべて引き続き遵守することを保証しなければならない。

2.5 取締役会、またはその委任を受けた委員会は、以下の責任を負う：

- a. AI リスクマネジメントの全体的なガバナンスアプローチを承認すること。これには、金融機関の AI リスクを継続的に評価・管理するための主要な枠組み、構造、方針、手順が含まれる。
- b. 重要性が認められる AI リスクが、金融機関のリスク許容度枠組み内で明示的に扱われることを確保すること。これには適切な定性的な記述や定量的な指標・制限の設定が含まれる。
- c. AI リスクマネジメントの監督に関する取締役会と上級管理職の役割と責任を明確に設定すること。
- d. AI に関する十分な理解を確保し、効果的な監督と検証を行うこと。
- e. AI 利用のリスクマネジメントに関する金融機関のアプローチ、リスク許容度枠組み、役割と責任、能力、文化が、新たな AI の発展や金融機関のリスクプロファイル・事業戦略の変化に対応できるよう、定期的に見直されることを確保すること。

2.6 上級管理職は以下について責任を負う：

- a. 金融機関全体の AI 関連リスクマネジメント方針・手順が、金融機関のリスク許容度と整合性を保ちながら効果的に実施されることを確保すること。

²⁷ 金融機関は、3.8 項から 3.11 項に定めるリスク重要性評価に関するガイダンスに基づき、AI ユースケースの重要性と AI リスクへの全体的なエクスポージャーを評価すべきである。金融機関の AI リスクへの全体的なエクスポージャーの重要性は、3.10 項に定める次元に加え、金融機関の事業戦略及び全体的なリスクプロファイルへの影響を考慮すべきである。一般的に、金融機関において、当該金融機関またはその顧客に悪影響を及ぼす可能性のある重大なリスク（例：重要業務ラインや機能領域における高リスク AI の展開、規制対象業務の実施に関連する AI の使用など）に晒される AI ユースケースが一つ以上存在する場合は、当該金融機関は、そのような強化された管理監督、すなわち専任のクロスファンクショナル委員会の設置が必要かどうかを評価すべきである。

- b. AI 関連リスクマネジメント方針・手順の有効性を定期的に見直し、新たな AI 技術の発展、金融機関のリスクプロファイルや事業戦略の変化、関連規制要件に適合するよう適切に改訂すること。
- c. 金融機関全体で、AI 関連リスクマネジメントの調整と説明責任のための強固な仕組みを確立し維持すること。
- d. 重大な AI リスクや例外事象（インシデントやリスク閾値超過など）を管理するための内部エスカレーションプロセスを確立し、適切かつ迅速な対応が取られることを確保すること；
- e. 重要な AI リスク問題について、取締役会にタイムリーに報告すること。
- f. 効果的な AI リスクマネジメントのために必要な人材の能力を確保し、適切な研修や能力構築を含む、人的・技術的・財務的資源などの十分なリソースを配分すること。

3 主要な AI リスクマネジメントシステム、方針及び手順

- 3.1 **金融機関は、AI リスクマネジメント枠組みが、AI の識別、棚卸し、およびリスクの重要性評価のための主要なシステム、方針、手順を網羅していることを確保すべきである。**金融機関は、AI の識別、棚卸し、およびリスクの重要性の判断に一貫性のある堅牢なアプローチを適用し、評価されたリスクの重要性に比例した統制を適用すべきである。

AI の特定

- 3.2 **金融機関は、関連する全ての事業領域及び機能領域における AI の使用を一貫して特定するためのシステム、方針及び手順を確立すべきである。**この特定プロセスは、AI のライフサイクル全体を通じて適切なガバナンス、リスクマネジメント標準及び統制を効果的に適用するための重要な前提条件である。この特定プロセスを円滑に進めるため、堅牢なシステムに支えられた明確な定義、規準及びプロセスを実施すべきである。
- 3.3 **金融機関は、AI の識別に関する明確な役割と責任を割り当てるべきである。**これには、金融機関全体での識別プロセスの一貫した適用確保、証明プロセスの設定、AI の使用の有無を最終的に判断する仲裁者としての役割など、**AI 識別システムおよびプロセスを担当する統制機能の指定が含まれる。**指定された統制機能は、識別プロセスと結果の明確な文書化が維持されること、また識別システムとプロセスが定期的に見直され更新され、新しい AI 技術を考慮に入れることも確保すべきである。

AI インベントリ

- 3.4 **金融機関は、AI のライフサイクル全体におけるガバナンス、監督、リスクマネジメントを支援するため、金融機関全体における AI のユースケース、システム、モデルについて、正確かつ最新のインベントリを作成・維持すべきである。**インベントリの保守に関する明確な方針と手順を定め、新規、更新、廃止された AI ユースケース、システム、モデルを正確に反映させる必要がある。金融機関は、既存のインベントリを拡張して AI ユースケース、システム、モデルを含めるか、AI ユースケース専用のインベントリを確立することができる。いずれの場合も、AI インベントリと金融機関内の他の関連インベントリとの間に明確な連携を確保すべきである。
- 3.5 **AI インベントリは、効果的なガバナンスと監視、ならびにリスクマネジメントを可能にする主要な属性を把握すべきである。**具体的な属性は金融機関の状況によって異なるが、各 AI ユースケース、システム、モデルについて以下を含む可能性がある：目的と説明、承認された使用範囲（例：管轄区域）、モデルタイプ、使用データ、依存関係、ライフサイクルステータス、割り当てられたリスク重要度評価、妥当性確認

テータス、主要な役割と責任（例：所有者、開発者）、および必須文書へのリンク。この情報を維持することで、リスクの集約、適用範囲の順守状況の監視、および統制の一貫した適用が容易になる。

- 3.6 **インベントリの設計は定期的に見直し、捕捉する属性が、追加の関連属性やガードレール、あるいはサードパーティの AI に関する追加情報が必要となる可能性のある新しい AI 技術を考慮したものとなるよう確保すべきである。**
- 3.7 **金融機関は、AI のインベントリ作成に関する明確な役割と責任を割り当てるべきである。**これには、インベントリに関する方針や手順、インベントリの保守・更新、認証プロセス、インベントリ範囲の定期的な見直しなど、**AI インベントリを担当する統制機能の指定が含まれる。**

AI リスク重要性アセスメント

- 3.8 **金融機関は、その事業の性質に基づき、AI のユースケース、システム、またはモデルのリスク重要性を評価するためのアセスメント手法を確立すべきである。**この評価手法は、金融機関が使用する各 AI ユースケース、システム、モデルに対して、一貫してリスク重要性アセスメントを実施するために適用されるべきである。こうしたアセスメントは、AI リスクマネジメントアプローチを調整し、AI がもたらすリスクに見合った統制が適用されることを確保するために重要である。これにより、リスク重要性が高いと評価された AI ユースケース、システム、モデルは、より厳格な精査と強固な統制を受け、効果的な監視を支援し、AI の使用が承認された範囲内にあることを確認できる。
- 3.9 **アセスメントでは、適切なリスクマネジメント統制を適用する前の AI ユースケース、システムまたはモデルに内在するリスクの重要性和、リスクマネジメント統制適用後の残存リスクの重要性の両方を考慮すべきである。**金融機関は、AI ユースケース、システムまたはモデルの残存リスクの重要性が、展開前に金融機関のリスク許容度に合致していることを確認すべきである。また、各 AI ユースケース、システム、モデルに対するリスク重要性評価の方法論および評価結果は、継続的な妥当性と適切性を確保するため、定期的に見直すべきである。
- 3.10 **リスク重要性評価では、金融機関の状況に関連する様々なリスク次元を考慮すべきであり、少なくとも以下を網羅する：**
 - a. **影響：**AI システムまたはモデルの故障、誤動作、性能不振が金融機関（例：財務的、業務的、規制的、評判的）および顧客やその他の利害関係者（例：公平性、倫理違反、消費者保護）に及ぼす潜在的な結果。AI システムまたはモデルが処理するデータの性質と機密性も考慮すべきである。
 - b. **複雑性：**使用される AI 技術の性質、その応用における新規性、または使用するデータから生じる。AI 技術への理解が時間とともに進化するにつれ、このリスク次元も変化しうる。例えば、研究の進展や理解の深化により、当初は十分に理解されていなかった新しい AI 技術の複雑性が変化する可能性がある。
 - c. **依存度：**AI システムまたはモデルに付与される自律性のレベル、それが支援するプロセスにおける人間の関与または監視の程度、代替手段の有無を考慮する。
- 3.11 **金融機関は、AI リスクの重要性評価について明確な役割と責任を割り当てるべきだ。**評価プロセスが金融機関全体で一貫して適用されるよう、統制機能を割り当て、証明プロセスを設定し、明確な文書化が維持されることを確保し、AI ユースケース、システム、モデルのリスク重要性を決定する最終的な判断者としての役割を果たすべきだ。

4 AI ライフサイクル²⁸ 管理

- 4.1 **金融機関は、AI のユースケース、システム、モデル全体のライフサイクルをカバーする堅牢な統制を計画・実施し、それらの統制に関する明確な役割と責任を割り当てるべきである。** AI ライフサイクル統制は、新しい AI 技術の利用を考慮して定期的に見直すべきである。
- 4.2 **各 AI ユースケース、システム、モデルについて、金融機関はユースケースを明確に定義し、異なる事業部門や機能部門にわたるライフサイクル全体における役割と責任を明確に割り当てるべきである。** 金融機関はまた、リスクの重要性アセスメントを実施し、入手可能な情報を AI インベントリに記録すべきである。こうしたアセスメントとインベントリ情報は、新たな情報が入手可能になった際に、AI ライフサイクル全体を通じて見直しと更新を行うべきである。
- 4.3 ²⁹**金融機関は、AI の重要性評価結果に適合するよう、関連する AI ライフサイクル管理措置を比例的な方法で実施できる。** サードパーティ AI プロバイダによる情報開示が不十分であるなど、実務上の制約がある場合、金融機関は当該制約から生じるリスクを識別し、AI の使用制限など必要な緩和を講じるべきである。こうした緩和を実施した後、金融機関は残存リスクが自社のリスク許容範囲内に収まることを確保すべきである。
- 4.4 **高リスクと評価された AI のユースケース、システム、モデルについては、金融機関は緊急時対応計画を策定し実施すべきである。** これらの計画は、AI の故障や予期せぬ動作が発生した場合の事業継続を確保するため、代替システムや手動プロセスなどの代替案を明記すべきである。計画は定期的に見直し、有効性をテストすべきである。「キルスイッチ」を備えた AI については、³⁰、明確な緊急時対応プロトコルを整備し、定期的なテストすべきである。

データ管理

- 4.5 **金融機関は、AI ライフサイクル全体で使用されるデータが目的に適合し、代表者であり、高品質で、堅牢なデータガバナンスの対象となるよう、データ管理統制を整備すべきである。** データ所有権、アクセス管理、知的財産権などの分野における一般的なデータガバナンスおよび管理標準は、AI に使用されるデータにも適用され、必要に応じて AI 固有の要件に対応するために強化されるべきである。金融機関が考慮すべき主な領域には以下が含まれる：
 - a. **目的適合性：** AI ユースケース、システム、モデルで使用されるデータの、意図された目的と文脈に基づく適合性。公平性などの他の考慮事項に対するデータ使用の評価を含む。
 - b. **データの代表者：** AI が使用されるストレス状態を含む、現実世界の全条件範囲において、AI ユースケース・システム・モデルの訓練およびテストに使用されるデータの代表者。

²⁸ AI ライフサイクルとは、AI の構想段階から廃止・廃棄に至るまでの進化過程を指す（ISO/IEC 22989「AI の概念と用語」より改変）。

²⁹ 特定のケースでは、AI は最初から完全に展開されず、パイロット運用や段階的展開といった部分的な展開となる場合がある。こうした AI の部分展開では、完全展開向けに設計されたライフサイクル標準、プロセス、統制の調整が必要となる場合がある。例えば、特定テストは部分展開からの情報が得られた後のみ実施される場合や、導入前レビューが完全に完了していない場合がある。このような部分的な展開に対しては、金融機関は標準的なライフサイクル基準、プロセス、統制からの逸脱を管理するための明確な方針と手順を確立すべきである。標準的なライフサイクル基準、プロセス、統制からの逸脱を管理するための明確な方針と手順には、時間制限やユーザー制限、成功の明確な規準、所有者およびエンドユーザーの利用条件、使用パターンや出力の異常に対する厳密な監視、限定された使用範囲の遵守の確保などが含まれる可能性がある。

³⁰ 「キルスイッチ」とは通常、リスク許容度を超えた場合に AI を迅速に無効化できる仕組みを指す。

- c. **データの品質**：AI ユースケース、システム、モデルで使用されるデータの品質の適切性。データの関連性、正確性、完全性、最新性のアセスメント、ならびにデータ品質の定期的な監視、異常値、ドリフト、潜在的なバイアスのチェックを含む。
- d. **データ格付**：AI ユースケース、システム、モデルにおけるデータの適切な利用を導くための適切なデータ格付プロセス。使用されるデータの重要度と機密性を考慮する。
- e. **データセキュリティ**：安全なデータライフサイクル管理の実践。これには、廃止または不要となったトレーニングデータ、モデル成果物、出力の適時な破棄または消去が含まれる。適切なデータ保護措置（転送中および保存時のデータ暗号化、安全なデータ取り扱いを含む）も含まれ、AI ユースケース、システム、モデルの入力と出力を保護する。
- f. **データ・プライバシー**：データ・プライバシーに関する関連規制要件およびガイダンス³¹）に基づく適切なデータ・プライバシー対策を実施する。また、AI モデルのトレーニングに顧客や従業員の機微な個人データを使用する場合、または AI がリアルタイムで当該データにアクセスすることを許可する場合、事前に許可を得る。
- g. **データの監査可能性とトレーサビリティ**：データソース、選択、処理、トレーサビリティといった主要なデータ管理側面の適切な文書化。データが目的適合性および代表者を有すると評価された方法。データに関連する承認および是正措置の記録。

透明性と説明可能性

- 4.6 ³²³³ **金融機関は、AI の使用事例、システム、モデルに求められる透明性と説明可能性の程度を、評価されたリスクの重要性に応じて決定し、それに依拠して関連する統制を確立すべきである。** 透明性は、AI の使用における説明責任と信頼を支える上で重要であり、顧客、内部ユーザー、その他の利害関係者が、AI の使用事例、システム、モデルのリスク、信頼性、限界について情報に基づいたアセスメントを行うことを可能にする。また、金融機関が AI の利用が表明した目的やリスクマネジメント標準に沿っていることを示すことを可能にし、顧客やその他の利害関係者間の信頼強化に寄与する。
- 4.7 **透明性と説明可能性の必要度に関する主な考慮事項には、最終決定における AI への依存度（すなわち AI の自律性の程度）、顧客またはリスクマネジメントの結果への影響レベルが含まれる。** 例えば、与信決定、保険引受、または顧客の結果に大きな影響を与えるその他の規制対象業務（例：金融アドバイザリーサービスの提供やファンド運用）で AI に大きく依存する場合、説明可能性に関するより厳格な標準が必要となる。こうした AI のユースケースでは、金融機関は入力データに含まれる様々な特徴や属性、それらの使用根拠、ユーザーが結果の主要な要因を識別する能力、顧客への AI 使用の告知の必要性、AI による決定の結果、救済手段について、より注意を払うべきである。

³¹ 例：個人情報保護委員会「AI 推薦・意思決定システムにおける個人データ利用に関する助言ガイドライン」
<https://www.pdpc.gov.sg/guidelines-and-consultation/2024/02/advisory-guidelines-on-use-of-personaldata-in-ai-recommendation-and-decision-systems>

³² 透明性とは、個人または集団に影響を与える AI の使用に関する開示、および関連性があり要求された場合の説明の提供を指す。説明可能性とは、AI が生成した出力や決定の理解を促進するための手法に関わる。説明可能性の要求水準は AI の用途によって異なり、要求される説明可能性の標準は、リスクの重要性や、AI による意思決定が説明を必要とする可能性（例：エンドユーザーの意思決定支援、金融機関の顧客への説明責任）を考慮に入れることがある。

³³ 例えば、NIST AI リスクマネジメント枠組み（<https://doi.org/10.6028/NIST.AI.100-1>）で扱われている透明性と説明可能性に関連する領域に基づく。

公平性

- 4.8 **金融機関は、自らが「公平」とみなす結果を定義し、AI ライフサイクル全体を通じて有害なバイアスや差別的な結果を識別・緩和するための適切な管理措置を講じるべきである。**その際、評価されたリスクの重要性に応じて調整を行う必要がある。例えば、信用決定や 保険引受などの分野で使用される AI は、個人に対する金融サービスや商品の不公平なアクセスや拒否につながる可能性があるため、より注意を払うべきである。
- 4.9 公平性の考慮が関連する場合、金融機関は公平性評価を実施すべきである。これには、関連する保護属性³⁴の定義、適切な公平性指標を用いた AI の使用が特定のグループに体系的な不利益をもたらすかどうかの評価、結果と講じた緩和の文書化が含まれる。

人間の監督

- 4.10 **金融機関は、AI のユースケース、システム、モデル³⁵のライフサイクル全体にわたって、適切な人的監視を確保するための統制を整備し、定期的に見直すべきである。**人的監視の必要性和程度は、AI 利用の目的を考慮し、利用される AI のリスク重要性に比例したものであるべきである。また、AI の利用が速度と規模を増すにつれて生じる自動化バイアスや意思決定疲労も考慮すべきである。金融機関が考慮すべき主な領域には以下が含まれる：
- 役割と責任：**人的監視に関する役割と責任を明確に割り当てる。これには人的監視に関連するエスカレーションおよび意思決定プロセスも含まれる。
 - 能力：**AI 利用を監視する担当者に必要な能力を付与すること。これには介入に必要な権限と能力を含む。
 - 設計：**適切な人的監視を可能にし促進するよう、AI システムやモデルを最初から設計・開発すること。
 - 文書化とレビュー：**人的監視の決定や介入（インシデントやニアミスを含む）を文書化し定期的にレビューするプロセスを確立し、人的監視の有効性を評価する。

サードパーティによる AI 管理

- 4.11 **金融機関は、サードパーティ AI を導入・開発・展開する際の管理措置が、サードパーティ AI を使用または依存するユースケース・システム・モデルのリスク重要性に十分対応していることを確保すべきである³⁶。**これには、金融機関のユースケース（自社のデータ使用を含む）の文脈でサードパーティ AI 製品・サービスをテストすること、およびサードパーティ AI プロバイダによる不十分な開示から生じる情報ギャップに対処するための補完的テストの実施が含まれる。金融機関はまた、サードパーティが開発した AI が適切なレビューの対象となることを確保するとともに、サードパーティ AI の更新や変更に関する通知を受け取るプロセスを確立し、そのような更新や変更の影響を管理・アセスメントすべきである。金融機関が考慮すべき主な領域は以下の通りである：

³⁴ 防御対象属性と代替関係にある属性、または防御対象属性と高い相関性を持つ属性についても考慮すべきである。

³⁵ これには、人間の関与がどのように実施されるか（例：人間がループ内またはループ外にいる）にかかわらず、金融機関が人間の関与が必要であると評価したすべての事例が含まれる。

³⁶ アウトソーシングや第三者サービスの利用に伴うリスクマネジメントに関する MAS の期待と要件は、第三者の AI の利用にも適用される。

- a. **透明性**：データ、モデル、技術、サイバーセキュリティリスクなど主要リスクが、当該サードパーティ AI の開発・展開過程でどのように対処されているかについて、サードパーティ AI プロバイダが提供する透明性のレベルをアセスメントすること。また、金融機関全体においてサードパーティ AI プロバイダに求められる透明性のレベルについて、明確かつ一貫した期待値を定めること。透明性と説明可能性が求められるにもかかわらず、サードパーティ AI においてそれが得られない場合、金融機関は、サードパーティ AI の挙動を理解するための追加テストの実施、人的監視の強化、利用者への適切な開示など、代償的措置の採用を検討すべきである。
- b. **公平性**：金融機関は、サードパーティ AI の利用による公平性の結果について責任を負うため、サードパーティ AI プロバイダの公平性に関する慣行について十分な注意を払うべきである。
- c. **サプライチェーンアセスメント**：主要なサードパーティおよびオープンソースの AI モデル、データセット、依存関係について、モデルの出所、トレーニングデータの完全性、その他の既知の脆弱性を含むサプライチェーンリスクアセスメントと妥当性確認が実施されているか確認すること。
- d. **集中リスク**：主要な第三者 AI プロバイダへの過度な依存（直接的・間接的を問わず）から生じる潜在的な集中リスクを評価する。
- e. **緊急時対応計画**：特にリスクの重大性が高いユースケース、システム、モデルで使用されるサードパーティ AI について、潜在的な障害、予期せぬ動作、ベンダーによるサポート終了に対処するための堅牢な緊急時対応計画を策定する。
- f. **法的契約**：法的契約を更新し、期待と責任をより明確にする。具体的には、性能保証、データ保護、監査権、AI 導入時の通知、AI 組み込み前の金融機関の同意取得に関する条項など。
- g. **能力**：サードパーティ AI の調達、開発、展開、利用に関わるスタッフの認識向上と能力開発を行う。
- h. **複雑性**：金融機関が経験の浅い複雑なサードパーティ AI 製品・サービスを利用する場合、より詳細なアセスメントを実施する。³⁷

選定

4.12 **使用する AI アルゴリズム³⁸ やデータの特徴量³⁹ を選定する際、金融機関は AI ユースケース・システム・モデルの目的とリスクを考慮すべきである。**特に、単純な代替案や従来型手法よりも複雑なアルゴリズムや理解度の低い特徴量を選択する場合、開発者に対し選定プロセスの正当性と文書化を求めるべきである。これには、ユースケースの要件や性能要求と、複雑性・公平性の必要性・透明性・説明可能性といった要素のバランス調整が含まれる場合がある。選択は可能な限り理論、研究、または業界で認められた慣行によって裏付けられるべきである。金融機関はまた、アルゴリズムや特徴量の選択が AI 利用の文脈と

³⁷ 例えば、金融機関がサードパーティのプロバイダが提供する AI エージェントの開発や展開の経験がない場合、より詳細な評価やテスト、技術やサイバーセキュリティのチェックが必要になる可能性がある。

³⁸ これには特定の AI モデル（例：特定の機械学習モデルや深層学習モデル）や技術（例：最適化や微調整技術）が含まれる可能性がある。

³⁹ 特徴量とは、データセット内のデータポイントの属性を指す。例えば、融資に関連するデータの場合、債務者の収入と融資の残高は、2つの可能な属性または特徴量である。特徴量エンジニアリングとは、AI モデルの性能向上のために、データセットの元の属性から特徴量を選択、修正、または新規作成するプロセスを指す。例えば、債務者の収入と貸出残高を 0 から 1 の範囲の共通尺度で正規化すること、あるいは既存の属性から債務収入比率などの新たな派生特徴量を作成することが挙げられる。

整合していることを確保するため、ドメイン専門家やユーザー（例：関連する事業部門や機能ユニットの専門家やユーザー）によるレビューの組み込みを検討すべきである。

- 4.13 理解が進んでいない、より新しく複雑な AI アルゴリズムが選択される場合、金融機関は、幻覚、不透明性、セキュリティリスク、および金融機関がそのようなリスクを緩和する能力など、金融機関に対する新たなリスクや高まったリスクと、そのような AI アルゴリズムの展開によるメリットを慎重に比較検討すべきである。

評価とテスト

- 4.14 **金融機関は、AI のユースケース、システム、モデルの評価されたリスクの重要性に見合った適切な評価とテストを実施すべきである。**各 AI ユースケース、システム、モデルは、展開前に評価されたリスクの重要性に基づき、適切な信頼性と安全性の水準を満たすよう評価およびテストされるべきである。金融機関は主要な AI リスクを識別し、明確かつ測定可能な閾値を設定すべきである。AI リスクの識別および信頼性・安全性アセスメントは、AI 利用の文脈を考慮すべきであり（ ）、アセスメントはベストプラクティスを参照すべきである⁴⁰。評価とテストは、現実世界のシナリオからエッジケースに至るまで、様々な想定可能な条件下における AI ユースケース、システムまたはモデルの性能を評価すべきである。金融機関が考慮すべき主要な領域には以下が含まれる：

- a. **評価指標**：AI の目的と整合した適切な評価指標を定義し、これらの指標に対する許容可能な性能閾値を設定する。性能閾値は明確に定義・文書化し、事業責任者、開発者、レビュー担当者間で合意を得る必要がある。
- b. **テスト手法**：アウトオブサンプル/アウトオブタイムテスト、感度分析、異なるデータ分布や期間における安定性分析、サブ集団分析、ストレステスト（必要に応じてエッジケースや敵対的テストを含む）、エラー分析、代替手法とのベンチマーク比較など、関連するテスト手法を採用する。テスト用データセットは、金融機関における使用環境を代表する者とする。
- c. **過学習の緩和**：可能な限り、特に複雑な AI に対して過学習を防ぐ技術を導入する。これには、明確な性能向上が正当化されない限り単純なモデルを優先すること、複雑さの制約（例：AI モデルの正則化⁴¹）、適切な特徴量選択、交差検証などの堅牢な妥当性確認技術の使用が含まれる。

- 4.15 **開発段階で AI に関連するリスクや制限が識別された場合、FI は展開前にこれらのリスクや制限を緩和するための適切な管理策とガードレールを設けるべきだ。**生成的 AI や AI エージェントなどの新しい AI 開発の評価とテストでは、その主要な故障モードを網羅すべきである⁴²。

⁴⁰ 例えば、IMDA の「LLM ベースアプリケーション安全テスト用スターターキット」は <https://www.imda.gov.sg//media/imda/files/about/emerging-tech-and-research/artificial-intelligence/large-language-model-starter-kit.pdf> でアクセス可能だ。このスターターキットは、LLM ベースアプリケーションのテストに関する新たなベストプラクティスや手法をまとめた自主的なガイドライン集である。

⁴¹ こうした手法は一般的に、使用するパラメータ数を制限して学習済みモデルの複雑さを抑えようとする。例えば、一部の正則化手法は重要度の低いパラメータをゼロに強制する。

⁴² 例えば、生成的 AI の評価とテストでは、幻覚現象、有害・偏った内容など望ましくないコンテンツの生成、バイアス、データ漏洩や開示、敵対的攻撃への脆弱性などに特に注意を払うべきだ。こうしたリスクのテストに関する詳細は、IMDA の「LLM ベースアプリケーションの安全性テストのためのスターターキット」を参照のこと。

技術・サイバーセキュリティリスク

4.16 ⁴³金融機関は、AI システムが安全で適切にガバナンスされ、技術リスク及びサイバーセキュリティリスクを管理するための適切な統制によって支えられていることを確保すべきである。技術リスクマネジメントに関連する規制要件及びガイダンスが適用される。金融機関が考慮すべき主要な領域には以下が含まれる：

- a. **セキュリティ**：展開は、強化された構成、ネットワークセグメンテーション、入力検証、API 認証、暗号化、データ損失防止などの技術的統制を含む、安全な IT 環境で行うべきだ。
- b. **アクセス管理**：展開中の AI コンポーネントおよびインフラへのアクセスは、役割ベースのアクセス制御と多要素認証により厳格に管理すべきである。特権アカウントのアクセス管理においては、この点が特に重要である。AI システムのトレーニングとテストを別々のチームに割り当てるなど、職務分離も検討すべきである。
- c. **サードパーティ**：プラグインや API などのサードパーティ製コンポーネントやサービスを使用する場合、使用を管理し、データエクスポージャーを制限し、セキュリティ、コンプライアンス、運用リスク（サイバーセキュリティ問題やサードパーティプロバイダのサービス中断など）を監視するための統制を実施すべきである。

再現性と監査可能性

4.17 金融機関は、再現性と監査可能性を確保するため、AI 開発プロセスを文書化すべきである。文書化は、レビュー担当者や監査人などの独立した第三者が、AI システムやモデルの実装とその結果を理解し、必要に応じて再現できるほど詳細である必要がある。文書化標準は開発プロセス全体を網羅し、以下のような情報を含む場合がある：

- a. データソース、処理、品質チェック；
- b. 選択の根拠；
- c. トレーニング手順（コードバージョン、環境、ハイパーパラメータを含む）；
- d. 評価指標と性能閾値、テスト手法、結果；
- e. 説明可能性分析、公平性アセスメント；および
- f. 主要な仮定、制限事項、および緩和策。

展開前レビュー

4.18 展開前に、金融機関は AI のユースケース、システム、またはモデルを、その開発に関与していない関係者によるレビューに付すべきである。これにより、評価やテストなどの関連する 管理が遵守されていることを確認する。こうしたレビューは明確な方針と手順によって支えられ、定義された範囲内で AI が意図された目的に対して適切性、堅牢性、性能を備えているかを評価すべきである。こうしたレビューに関連する役割と責任は明確に割り当てられ、展開前レビューの範囲と独立性は、AI の評価されたリスクの重要性に見合ったものでなければならない。

⁴³ 例えば、MAS の「技術リスクマネジメント実践ガイドライン」は

<https://www.mas.gov.sg/regulation/guidelines/technology-risk-management-guidelines> で閲覧可能だ。

- 4.19 これらの導入前レビューの範囲、および評価・レビューを実施する当事者の独立性の程度は、AI のユースケース、システム、またはモデルの評価されたリスクの重要性に見合ったものであるべきだ。リスクの重要性が高いと評価された AI のユースケース、システム、またはモデルは、導入前に正式な独立した妥当性確認を受けるべきだ。このような妥当性確認は、必要な専門知識と客観性を有し、開発チームや展開チームから独立した有能な担当者または機能によって実施されるべきだ。妥当性確認プロセスは開発者に対して効果的な検証を提供し、以下の領域をカバーすべきである：
- a. 設計の概念的妥当性；
 - b. データ入力の適切性と品質；
 - c. 実装の完全性；
 - d. 評価指標、性能閾値、テスト手法および結果；
 - e. 説明可能性分析と公平性アセスメント；および
 - f. 前提条件、制限事項、および緩和策。
- 4.20 AI のユースケースにおいて、高いリスク重要性があると評価されなかったシステムやモデルについては、開発・展開プロセスにおける主要な側面について、他の形式の文書化されたレビュー（開発や展開に関与していない有資格者によるピアレビューなど⁴⁴）を実施することができる。
- 4.21 **レビュー結果（識別された制限事項、必要な是正措置、使用条件を含む）は、レビュー担当者が関連する承認団体に報告すべきである。** 関連する承認団体は、レビューの推奨事項が適切に実行されることを確保すべきである。
- 4.22 **金融機関は、AI を管理された安全な方法で本番環境に展開できるよう、技術およびサイバーセキュリティレビューを実施すべきである。** こうしたレビューは、技術的実装、システムおよびネットワークセキュリティ、システムのレジリエンス、復旧可能性、展開に向けたAIの運用準備状況などの領域をカバーすべきである。金融機関が考慮すべき主要領域には以下が含まれる：
- a. **セキュア設計とアクセス管理：** API、プラグイン、サードパーティサービスを含む AI システムアーキテクチャについて、セキュア設計、環境分離、最小権限の原則に基づくアクセス管理を審査する。稼働前には、暗号化、データ損失防止、ファイアウォール、アクセス制限、ロギングなどの重要制御が動作し正しく設定されていることを確認するため、展開チェックリストを完了させる。
 - b. **テストと脅威軽減：** AI システム全体で脆弱性評価、侵入テスト、レッドチームを実施し、セキュリティ上の弱点を識別・修正する。敵対的脅威（操作、回避、汚染など）に対するAIシステムのテストを行い、入力妥当性確認、スロットリング、異常検知などの安全対策を導入する。

展開後の監視とレビュー

- 4.23 **金融機関は、展開済みの全 AI（金融機関で使用されるサードパーティ AI を含む）を継続的に監視するための包括的かつ堅牢な管理策を開発・実施すべきである。** AI に伴う不確実性、その動的な性質、時間の経過に伴うデータやモデルのドリフトによるモデルの陳腐化や性能低下の可能性を考慮すると、展開済みの AI が意図した通りに動作し、継続的に目的に適合していることを保証するためには、継続的な監視が不可欠である。監視活動の頻度と強度は、展開済み AI の評価されたリスクの重要性に見合ったものであるべきだ。金融機関が考慮すべき主な領域には以下が含まれる：

⁴⁴ これには、AI の初期展開に関与する個人と、AI システムの運用など継続的な展開に関与する個人の両方が含まれる。

- a. **監視措置**：AI ユースケース、システム、またはモデルのリスクに基づき、監視対象となる主要指標と各指標の許容性能閾値を定義する。指標には、頑健性、安定性、データ品質、公平性などの関連次元を含めることができる。劣化を未然に防ぎ早期対応を促すため、段階的な閾値（例：早期警戒閾値レベル）を検討すべきである。データドリフト（入力データ分布の変化）、コンセプトドリフト（入力と出力の関係の変化）、モデル全体のドリフトに対する適切なチェックを実施すべきである。関連する場合、推論プロセス、実行されたアクション、使用されたツールなど、AI を利用するワークフロー全体の情報フローと意思決定経路も監視すべきである。
 - b. **インシデントおよび問題管理**：監視プロセスから違反や異常が発生した場合、インシデントや問題の報告、追跡、エスカレーション、解決のための堅牢なプロセスを確立する。問題の深刻度や性質に応じて、AI モデルやシステムの再トレーニング、調整、再開発、廃止といった解決策が考えられる。高リスクかつ重要度の高い AI については、リスク許容度を超えた場合に AI システムを迅速に停止させる「キルスイッチ」やオーバーライド機構の導入を検討する。展開済み AI のユーザーに対し、継続的な改善、問題の是正・解決を支援するため、フィードバックの提供や問題報告を行う適切な手段をする。
 - c. **役割と責任**：展開済み AI の継続的監視とインシデント管理に関する明確な役割と責任を割り当てる。これには適切な責任者の指定を含む。
 - d. **文書化**：監査可能性と継続的なリスクマネジメントのため、監視活動、結果、識別された問題やインシデント、およびその後の是正措置について明確な記録を維持する。AI モデルやシステム、トレーニングデータ、パイプライン、設定ファイルへのアクセスは厳格に管理し、ログを記録すべきである。
 - e. **訓練と認識**：効果的な監視とインシデント管理を支援するため、監視責任者に適切な訓練を施す。また、展開済み AI の利用者には、意図しない動作を報告するために必要な訓練と認識を付与する。
- 4.24 **金融機関は、全ての AI 利用事例、システム、モデルにわたる総合的なリスクを定期的に見直すべきである。必要に応じて、ポートフォリオ内の特定の AI について、より詳細な見直しや再妥当性確認を実施する。リスクの重要性が高いと評価された AI 利用事例、システム、モデルについては、独立した第三者による定期的な再妥当性確認を実施すべきである。**こうした見直しや妥当性確認は、AI が適切に利用され続け、リスクが適切に管理されていることを保証するものである。特定の AI に対する詳細な妥当性確認や見直しの必要性は、その AI の評価されたリスク重要性に基づいて判断されるべきである。また、展開済み全 AI の総合リスクに関する定期的な見直し結果、継続的モニタリングによるアラートや違反、AI またはその運用環境における重大な変更、規制や技術の発展など外部環境における新たなリスクの特定によっても、こうした見直しが引き起こされることがある。

変更管理

- 4.25 **金融機関は、展開済み AI の変更を管理するための包括的かつ強固な統制を策定し実施すべきである。**こうした統制は、変更が意図しない動作、性能低下、展開済み AI の意図した用途との不整合を引き起こさないことを確保するために不可欠である。金融機関が考慮すべき主な領域には以下が含まれる：
- a. **変更の範囲**：重要な変更（例：モデルやシステムアーキテクチャ、主要な前提条件、意図された用途の変更）と軽微な変更（例：更新されたデータによる再トレーニング）を明確に定義する。重要な変更は、実装前に適切なレビューと再承認プロセスをトリガーすべきである。
 - b. **変更管理**：AI システムやモデルへの不正な改変を防止し、AI の変更（基盤コード、設定、トレーニングデータなど）がサービス停止、運用障害、モデル性能の低下といった意図しない結果を招

かないよう、変更管理メカニズム（例：ヒューマン・イン・ザ・ループ）を導入する。これには、AI モデルコードだけでなく関連データ、パラメータ、ハイパーパラメータ、その他の主要な成果物への変更を追跡し、トレーサビリティ、監査可能性、および以前のバージョンへのロールバック機能をサポートするバージョン管理システムが活用できる。

- c. **動的更新**：自動更新を可能とする厳格な正当性判断、自動更新対象の明確な定義（再トレーニングやハイパーパラメータ変更は許可するがコアアーキテクチャ変更は許可しないなど）、強化されたデータ品質チェック、より厳格な性能監視など、自動更新を前提に設計された AI 向けの強化された制御を実施する。

4.26 **金融機関は、AI が不要になった場合やリスク許容度を越えた場合に備え、その最終的な廃止または停止に関する明確な管理策を開発・実施すべきである。**これらの管理策は、依存関係、データ保持ポリシー、本番環境からの安全な削除、適切な関係者への通知を考慮すべきである。

5 AI 能力とキャパシティ

AI リスクマネジメント能力

- 5.1 **金融機関は、AI ユースケース、システム、モデルの開発および展開に携わる要員に必要な能力と適切な行動を確保すべきである。**これには、必要な人材の採用、従業員が AI を適切に活用し、効果的な AI リスクマネジメントを行うためのスキル、知識、文化を身につけるための適切な研修と支援の提供が含まれる。AI ユースケース、システム、モデルのリスクプロファイルに見合った、人的、技術的、財務的資源を含む十分なリソースを割り当てるべきである。
- 5.2 **AI ユースケース、システム、モデルの開発、展開、維持に関わる関係者が、効果的な AI リスクマネジメントのための十分な能力とキャパシティを備えていることを確認するため、定期的な見直しを行うべきである。**これには、効果的な AI リスクマネジメントのための能力とキャパシティ構築プログラムを定期的に見直し、最新のものに保ち、新しい AI 技術に関連するリスクに対処するための研修を取り入れることが含まれる。

AI のための技術インフラ

- 5.3 **金融機関は、自社の技術インフラが AI ユースケース、システム、モデルに十分対応できることを確保すべきである。**他の情報技術システムと同様に、AI システムおよびそれらで使用されるモデルも、システムの可用性、レジリエンス、安全性、サイバーセキュリティリスクを含む技術リスクに晒される。これらのリスクに対処するため、金融機関は、グラフィックス処理装置、ネットワークインフラ、システムメモリ、安全なデータパイプラインなどの基盤となるハードウェアおよびソフトウェアリソースが、AI のユースケース、システム、またはモデルに必要なパフォーマンス、スケーラビリティ、レジリエンスを満たすのに十分であることを確保すべきである。特定の AI ユースケース、システム、またはモデルに適した技術インフラを決定する際、金融機関は、関連する技術リスクマネジメントガイドラインおよび通知⁽⁴⁵⁾ ならびに関連業界枠組み⁽⁴⁶⁾ を考慮すべきである。

⁴⁵ <https://www.mas.gov.sg/regulation/guidelines/technology-リスクマネジメント-guidelines>

⁴⁶ 例えば、NIST AI リスクマネジメント枠組み (<https://doi.org/10.6028/NIST.AI.100-1>) で扱われる技術インフラ関連領域などである。

AI リスクマネジメントガイドラインの比例原則に基づく適用

1. 本ガイドラインは、規模やリスクプロファイルが異なる金融機関に対して比例的な適用を意図している。金融機関は、ガイドラインの適用可能性を評価する際、以下のフローチャートを参照することができる。金融機関における AI の利用が進化するにつれ、金融機関はガイドラインの適用可能性を再評価し、AI の利用拡大と責任あるイノベーションを支援するため、自らの枠組み、方針、プロセスを調整すべきである。

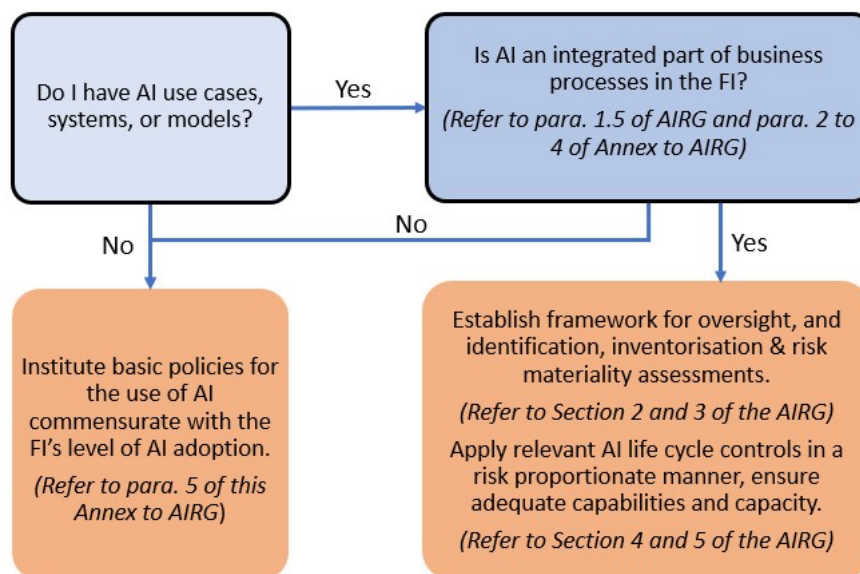


図 2 : AIRG の適用可能性の評価

2. 以下の指針となる質問のいずれかに「はい」と回答する場合、AI は金融機関の業務プロセスに統合された形で使用されていると見なされる。
 - a. AI サービスやツールへのアクセスが欠如した場合、金融機関が事業活動において実質的に依存している業務フローに支障をきたすか？
 - b. AI は、金融機関の事業活動において実質的に依存しているシステムと統合されているか？
3. ⁴⁷以下のいずれかの事例に類似した AI ユースケースを有する金融機関は、AI を業務プロセスの統合された一部として使用しているとは見なされないが、本附属書第 5 項に定める基本方針を適用すべきである：
 - a. 顧客へのメール返信のドラフトを支援するため、個々のアナリストがサードパーティ製大規模言語モデル（LLM）を利用する場合；
 - b. 顧客対応担当者は、顧客へのメール文法・スペルチェックや言い換えに AI ツールを活用する。
 - c. 投資調査チームのメンバーが、投資調査レポートを理解するために要約する AI を活用している；
 - d. 財務チームメンバーが Excel や可視化ツールで数式やグラフを生成する際に AI を活用する；

⁴⁷ 関与する人間は、これらのユースケースにおいて補助的な形で AI を利用し、出力結果を実際に使用する前に確認・検証を行うべきである。

- e. マーケティングスタッフが AI 画像生成ツールを活用し、マーケティング資料のデザインを支援する場合；あるいは
 - f. 保険金請求審査担当者が、請求書類の初期審査を支援するために AI を臨機応変に活用する場合。
4. 以下のいずれかの事例に類似した AI 活用ケースを持つ金融機関は、AI を業務プロセスの統合された一部として活用しているとみなされる：⁴⁸：
- a. 社内法務チームが標準契約書の AI 契約書レビューツールを使用している場合 - 法務チーム全体が体系的に使用しており、廃止すると契約書レビュープロセスが大幅に遅延する；
 - b. 生成的 AI を活用した社内 IT ヘルプデスクチャットボットを従業員の大半が利用 - IT サポートの第一窓口であり、これがなければヘルプデスクはチケット量に圧倒される。c. アナリストの大半が利用する PDF からの AI による財務データ抽出 - 大量の文書进行处理し、手動抽出では時間がかかりすぎる。
 - d. AI を活用したリサーチ集約ツール。投資アナリストの大半が使用。リサーチワークフローに組み込まれている。利用不可の場合、リサーチプロセスが大幅に遅延する。
 - e. AI 搭載投資アイデアスクリーナー：研究者の相当数が利用。アイデア創出プロセスに統合。廃止すればアナリストは手動で膨大な対象範囲をスクリーニングする必要が生じる。
 - f. 保険金請求審査プロセスへの AI 導入 - 補足書類进行处理するワークフローの中核を AI が担う；AI が利用不可の場合、請求処理が遅延する；または
 - g. 金融アドバイザープロセスへの AI 導入 - AI 搭載金融アドバイザーツールは顧客への金融助言や商品提案のワークフロー中核として使用される。利用不可の場合、助言・提案提供能力が大幅に低下または減少する。手動プロセスへの回帰は時間を要する。
5. AI 利用の基本方針は、金融機関の AI 導入レベルに見合ったものであるべきだ。こうした方針では以下の点を扱う必要がある：
- a. AI 監督責任者として上級管理職を指名すること；
 - b. AI の使用許可・禁止事項（例：機密情報、専有情報、顧客情報を公開 AI ツールに入力することの禁止）、使用前の全 AI 出力に対する人的レビューと妥当性確認の要件、承認済み AI ツールのリスト管理、新規 AI ツールの承認申請プロセスなどのガイドライン；
 - c. 全従業員へのガイドライン周知；
 - d. ガイドライン遵守状況の定期的な点検；および
 - e. ガイドラインの年次見直しと更新。
6. 金融機関は、本附属書で扱われていないガイドラインの比例原則に基づく適用に関連する事項について、金融管理局（MAS）に説明を求めることができる。

⁴⁸ 疑義を避けるために言えば、重大なリスクをもたらす全ての AI ユースケース（例：重要業務ラインや機能領域への AI 展開、規制対象業務の実施に関連する AI の体系的使用）は、ビジネスプロセスの統合された一部として AI を利用しているとみなされる。