「G7 サイバー・エキスパート・グループによる AI 及びサイバーセキュリティに関するステートメント」の仮訳

サイバー・エキスパート・グループ(CEG)は、G7 の財務大臣及び中央銀行総裁に、金融システムの安全性と強靭性にとって重要なサイバーセキュリティ政策上の問題について助言している。CEG は、生成 AI(GenAI)、AI エージェント（agentic AI）及びその他の先進的なシステムを含む人工知能(AI)技術の急速な発展を認識し、法域が進行中の変化をモニターし、官民学連携を促進し、AI がもたらす可能性のある新しく進展しているサイバーセキュリティ・リスクに積極的に対応することを奨励している。

（AI とサイバーセキュリティ：金融システムにおけるリスクと強靭性のナビゲーション）

AI は、金融セクターを変革する可能性のある一連の重要なイノベーションの最新のものである。AI 機能は長い間、複雑さ、統合、高度化の度合いを増しながらセクター全体で使用されてきた。これらの技術は、金融機関や当局のオペレーション能力、分析能力、リスク管理能力を向上させる大きな機会を提供する。

しかし、悪意ある主体が AI を取り込むことは、悪意あるサイバー活動の頻度と影響を増大させうる。また、AI システム、特に生成 AI 及び AI エージェントの複雑さと自律性の増大は、新たなサイバーセキュリティ・リスクを生じさせる。

本ステートメントは、ガイダンスや規制上の期待を示すものではない。むしろ、本ステートメントは、AI のサイバーセキュリティ上の特性について認識を高めることを目的とし、金融機関、規制当局、及び金融セクターにおいてセキュリティと強靭性を支えるその他のステークホルダーのための主要な検討事項を概説するものである。CEG は、金融当局が、金融機関、AI 開発者、テクノロジー企業、学術研究者、及びその他のステークホルダーと緊密に協力して、AI に関連するサイバーセキュリティの問題についての共通の理解を促進し、イノベーションを受け入れながらサイバー・リスクを軽減する戦略を策定することを奨励している。CEG は、引き続き、G7 の法域全体で対話を進めていく。

本ステートメントは、G7 の「基礎的要素」シリーズと併せて読まれるべきである。このシリーズは、サイバーセキュリティに不可欠なリスク管理の決定に関する内部及び外部の議論の指針となり、実効的なサイバー・リスク管理の実践を推進するために、法域及びセクターを越えた対話を促す。

（AI のサイバーインパクトの説明）

AI はサイバーセキュリティの状況を大きく変えている。以下の例は、AI がサイバー防御を強化し、既存の脅威を増幅し、AI システムの設計とデータ利用に起因する脆弱性をさらす可能性を強調している。

(1) AI がサイバー・レジリエンスを強化すること

AI は、膨大なデータセットを取り込み、変換し、学習することでサイバーセキュリティ業務を強化し、気づきにくいパターンや異常を特定し、対応時間を短縮することができる。

- **異常の検知と対応**：AI は、わずかなネットワークの異常を識別し、AI 搭載のウェブ・アプリケーション・ファイアウォール(WAF)などの適応型防御を強化し、機械のように高速でリアルタイムに脅威への対処ができる。
- **詐欺の検知と対応**：AI は、進化する支払いにおける詐欺のパターンを検知し、顧客の本人確認（Know Your Customer）の回避に使用されるディープフェイクを特定し、AI が生成したフィッシングメールにフラグを立てることができる。
- **予見的なメンテナンスとパッチ適用**：AI はシステム障害を予測し、ソフトウェアの脆弱性を検知し、脆弱性に対するパッチ適用に優先順位付けすることができる。
- **SOC の効率性**：生成 AI は、インシデントを要約し、対応を推奨することによって、セキュリティ・オペレーション・センター(SOC)を支援できる。
- **ベンダー及びサプライチェーンのリスク・モニタリング**：AI ツールは、財務指標及び公開データを用いてサードパーティ・リスクを分析できる。

(2) AI が既存のサイバー・リスクを増幅しうること

AI は、攻撃者がより正確、迅速、かつ大規模に攻撃することを可能にする。

- **AI を利用したフィッシングまたはなりすまし**：生成 AI 及び AI エージェントは、巧妙にパーソナライズされたフィッシング・メッセージやディープフェイクを生成する可能性があり、検知作業を複雑化させる。
- **自動化されたエクスプロイトの開発**：強化学習技術は、ネットワーク、ソフトウェアのパッケージ及びライブラリを検索することによって攻撃者を支援できる。偵

察をより迅速に実行することができるため、攻撃の効率向上につながる可能性がある。

- **マルウェアの開発と回避**：AI は、検知を回避するためにリアルタイムで進化するマルウェアを作成し、サイバー犯罪の敷居を引き下げ、攻撃の量と巧妙さの両方を高めることができる可能性がある。

(3) AI の脆弱性を狙った攻撃によるリスク

AI 自体が、サイバー攻撃の直接的なターゲットまたは脅威が発生する経路となる可能性がある。

- **データ・ポイズニング**：訓練中及び本番環境の両方で不正に操作されたデータは、モデルのパフォーマンス低下や、攻撃者が後で悪用できる隠れた脆弱性の埋め込みにつながる可能性がある。
- **データ漏洩**：公開されている AI ツールとのやり取りは、機密データの意図せざる開示または抽出につながる可能性がある。
- **プロンプト・インジェクション**：攻撃者は、システムプロンプトを不正利用して出力を操作できるほか、機密情報を取得できる可能性がある。

（安全性、健全性、コンプライアンス及び監督への影響）

AI がサイバー・リスクを軽減及び増幅する可能性は、規制対象企業及び監督当局に直接的に影響を及ぼす。

- **オペレーショナル・リスクとレジリエンス**：敵対的 AI は、システム停止、データ侵害、詐欺へのエクスポージャーを増大させる可能性がある。
- **人による監視**：人による監視が弱いと、インシデントの検知または対応が遅れる可能性がある。
- **モデル・リスク**：訓練または管理が不十分な AI モデルは、予測不可能な行動をとり、時間の経過とともに劣化する可能性がある。
- **サプライチェーン・リスク**：AI システムは、多くの場合、サードパーティのライブラリ、データセット、またはクラウドサービスに依存している。これらが侵害されると、サイバーセキュリティ防御の仕組みにバックドアまたは脆弱性が埋め込まれ、相互接続されたシステム全体でリスクが増幅される可能性がある。

- **AI リテラシー**：組織的な専門知識の欠如は、効果的なデプロイと監視を損なう可能性がある。

（リスクを管理しながら利用機会を最大化すること）

金融機関及び当局は、以下の方法により、安全な AI の導入に対する積極的なアプローチから便益を得られる可能性がある。

- サイバー・リスクと非サイバー・リスクの両方を管理し対処するために、AI をどこで、どのように利用できるかを特定する。
- 安全で責任ある AI、特に防御的用途のアプリケーションの開発に投資する。
- 目的に適合し、適切にリスク管理され、頑健な AI システムを推進する。
- 知識の共有と訓練を通じた社員の能力開発を支援する。

（金融セクターの検討事項）

AI は、AI 導入の経緯と成熟度に応じて、いくつかの方法で金融セクターにおけるサイバーセキュリティ関連リスクに影響を及ぼす可能性がある。

- **悪用**：悪意のある主体による AI の取り込みは、サイバーの脆弱性の悪用を促進し、高速化し、拡大させやすくする可能性がある。
- **サードパーティへの依存とサービス・プロバイダへの集中**：広く利用されている AI プロバイダにおける重大なサイバーインシデントは、多くの金融機関に影響を及ぼす可能性がある。金融セクターへの影響とシステム全体の障害の可能性は、AI サービスの代替可能性を含め、利用される AI サービスの性質と重要性に左右される。
- **専門的知見及び専門人材のリソース**：AI の専門的知見が不足している機関は、関連するサイバーセキュリティ・リスクに対して不釣り合いに脆弱である可能性がある。

（金融機関と金融当局の主要な検討事項）

AI 関連のサイバー・リスクを管理するために、金融機関は以下の質問を検討し得る。

- **戦略、ガバナンス及び監視**：ガバナンスの枠組みは新たな AI リスクに対応し

ているか。

- **サイバーセキュリティの統合**：AI システムは、セキュア・バイ・デザイン（企画・設計段階からのセキュリティ対策）の原則と整合しているか。
- **データの安全性と履歴管理**：データソースは精査され、履歴は管理されているか。
- **ログの取得及び監視**：異常及びエッジケースが記録され、レビューされているか。
- **アイデンティティと認証**：システムはなりすましや AI を利用した詐欺に対して耐性があるか。
- **インシデント対応**：インシデント対応計画と手順書は、AI により強化された攻撃と AI に特有のインシデントを考慮して更新されているか。
- **人的リソース、スキル、及び認識向上**：AI 利用を評価及び監視するための十分な専門的知見を確保する道筋は何か。

金融当局は、AI 及びサイバーセキュリティ・リスクに対処するために、以下の戦略を検討し得る。

- AI に特有のサイバーセキュリティ・リスクを理解するために組織内部の能力を強化する。
- AI とサイバーセキュリティに関連する強力なガバナンス及びリーダーシップへの関与を奨励する。
- テクノロジー企業、学界、金融業界のパートナーと協力して、進化する AI 能力を継続的に把握し、利用機会とリスクについて議論する。
- AI 関連リスクを既存のリスク管理プロセスに統合する。

効果的にデプロイされれば、AI は、検知と対応の速度と正確性を向上させ、システムの隠れた脆弱性を明らかにすることによって、サイバー攻撃耐性を強化できる。

（次のステップ）

AI がソフトウェア・システムや金融業務に深く組み込まれるにつれて、サイバーセキュリティへの影響は変化し続けうる。CEG は、金融セクターのステークホルダーに対し、以下を推奨する。

- サイバー防御能力の向上のために AI の可能性を探究する。
- AI に特有のサイバーセキュリティの脆弱性及びその緩和策を考慮してリスク・フレームワークをアップデートする。
- テクノロジー企業や学界と協力して共同研究と政策立案を行う。
- 金融セクターにおいて安全で信頼できる AI を推進するため、官民対話を推進する。

リスクを踏まえた慎重なアプローチにより、AI はサイバーセキュリティとレジリエンスのための効果的なツールとなり得ると同時に、金融システムの健全性と安定性の維持にも役立つ。