資料２－３

総務省

フロンティアAI時代の通信事業における
データの保護・セキュリティの在り方に関する検討会

フロンティアAI時代の
セキュリティ対策と課題

LINEヤフー株式会社

2026年11月　説明資料

本日の説明

全体方針から具体策、対策強化に向けた課題と対応、政府への要望

01 AIセキュリティの全体方針

AIの安全な利用と防御能力の強化を整理する共通の枠組み

02 AIの安全な利用と防御能力を高める取り組み

AI利用の制御、脆弱性の診断・修復、防御・復旧能力の強化

03 対策強化に向けた課題と対応

脆弱性対応の迅速化、AI修復・開発者支援、リスク管理の成熟度向上

04 政府に検討をお願いしたいこと

実際のリスクに応じた対応と、迅速な対策を支える制度・支援

AIセキュリティの全体像

AIの安全な利用と、攻撃の高度化に対応する防御能力を両面から整理

管理の観点	AIの安全な利用 Security for AI	防御能力の強化 AI for Security
統治・管理	責任・利用ルール・リスク判断	対策の優先順位・防御効果の評価
資産・リスク把握	AI・接続先・データ・権限の把握	攻撃面の把握・AIによる診断
防御	入力・外部連携・実行環境の制御	認証・権限管理・修復の高度化
監視・検知	AIの利用・操作ログ、異常の監視	AIによる分析・検知の高度化
インシデント対応	操作追跡・停止・アクセス失効	初動判断支援・封じ込めの自動化
復旧	設定・データの復元、再開判断	復旧手順の標準化・復旧能力強化

この表は対策領域を示し、全社での実装完了・有効性を示すものではない。

AIの入力・外部連携・実行環境と権限を管理する

AIの情報アクセスと操作に伴うリスクを、複数の制御点で抑える

制御点	主な取り組み・強化方針
入力の検査・標準設定	不正な指示の混入や個人情報を検知し、入力を検査・保護 全社員端末へのセキュリティ設定の配布
外部連携の検査	MCPサーバー（AIと外部ツールの接続）、スキル・プラグインを検査し、悪意ある動作や不適切な接続のリスクを確認
実行環境・権限の制御	AIエージェント用の隔離環境（Sandbox）の提供 エージェントを識別し、アクセス先・操作を必要最小限に制限

既存の取り組みと今後の強化方針を示す。適用対象・実装状況・有効性をそれぞれ確認する。

脆弱性の集約・検出・修復を高度化する

脆弱性を一元管理し、開発者とセキュリティエンジニアのAI活用で対応を強化

取り組み	主な内容
脆弱性の一元管理・優先順位付け	セキュリティ評価や検査ツールで検知した自社コード・OSSの脆弱性を一元集約し、管理対象を拡大 脅威情報を活用し、対応の優先順位を判断
開発者による脆弱性診断・修復の高度化	フロンティアAIのセキュリティ向けプログラムを活用し、診断・修復を高度化 セキュリティ組織と連携し、開発者自身が迅速に修復できる状態を目指す（詳細はP7）
修復の自動化・継続化	対応基準の見直しに伴う負担増を踏まえ、修復エージェントの展開を進め、修復を自動化・継続化して開発現場の対応力を高める（詳細はP7）
セキュリティエンジニアによるAI診断の高度化・効率化	高性能AIでソースコード・アプリを診断する仕組みを構築 診断を高度化・効率化し、AIをレッドチーミングにも活用 弱点を先制的に検知し、改善につなげる体制を整備

自動修復・開発者による迅速な修復は目指す状態。修正の検証・適用判断を組み込む。

高度化する攻撃への防御・復旧能力を強化する

AIの安全な利用と防御能力の強化を支える、共通基盤の高度化方針

強化する領域	目指す方向
ID・アクセスの保護 Identity Security / Zero Trust	人・システム・AIエージェントを識別し、権限を最小化 アクセスの継続的な検証と、不正利用の検知・対処を高度化
本番運用の安全性 Zero Touch Production	本番環境への直接操作を減らし、承認・検証・監査を 組み込んだ自動化によって、運用の安全性を高める
検知・初動対応	AIによるログ・アラート分析と、初動判断支援を展開 重要な判断への人の関与を含め、対応の自動化を進める
被害の抑制・復旧	侵害時の被害拡大を抑えるアクセス・通信制御を強化 バックアップの保護と復旧検証を通じ、復旧能力を高める

各施策は構想・検討を含む強化方針。AIの活用に加え、既存の防御・復旧対策も組み合わせる。

AIによる脆弱性修復の支援と自動化

増大する脆弱性対応の負担を軽減するため、開発者支援とOSS修復の自動化を進める

開発者の修復をAIで支援

フロンティアAI セキュリティ向けプログラムを開発者の脆弱性修復にも活用

脆弱性管理ツールで脆弱性とコード修復を
深刻度に応じて管理し、対応状況を可視化

OSS修復をエージェントで自動化

エージェントを安全に動かす環境を整備

修復エージェントを定期実行
コード修復とPR作成を自動化

人が修正内容を確認・判断してマージし、
その後デプロイ

脆弱性修復の効率化・自動化の仕組みを開発者に提供し、継続的に改善中

PR：変更の取り込み依頼。マージ：変更を本体のコードに取り込むこと。

AI脅威への対策を評価し、改善計画につなげる

これまでの対応を基盤に、共通の基準で対策を評価し、改善計画を具体化する

1

把握した課題に順次対応

個別の課題に応じた対策

把握したリスクや課題に応じて施策を実施

必要な対策を積み重ね、継続的に改善

2

体系的な評価

共通の基準で現状を評価

社内の成熟度フレームワークをAI脅威への対策評価に活用

対策状況と成熟度を評価し、現状のリスクと改善点を整理

3

改善計画を策定

評価結果を実行計画へ

リスクに応じて優先順位を設定
対策・担当・期限を具体化

改善計画を一元管理し、進捗と対策の効果を確認

評価で明らかになった課題を改善計画へ反映し、実施結果を次の評価につなげる。

対策の成熟度とリスクの大きさを区別して評価する。

先行して対策する事業者を後押しする制度へ

論点	検討をお願いしたいこと
防御的停止の考え方の整理	被害拡大防止に必要なサービス停止を事業者が判断できる環境・制度整備
実際のリスクに応じた評価	仮名化IDや暗号化等、情報漏えい時に悪用を困難にする対策の評価見直し 具体的な危険度や通知等の実効性に応じた報告・通知等制度の再設計
サプライチェーンにおける責任分担・支援	商用ソフトウェア供給者の責任明確化とOSSの修正・検証を支援 外部連携先やSaaS等に対する侵害の場合の各種対応の効率化の在り方の検討
情報共有の促進・啓発	被害企業による調査・開示・脅威情報の共有を促す仕組みの設計 フィッシング等、利用者側の被害を防ぐ啓発を強化
経済安全保障推進法との調和	経済安全保障推進法において、一定の安全性要件を満たすセキュリティ設備に対する審査期間短縮や特例運用（事後届出等）の検討

技術対策の強化と制度の実効性を両立し、利用者とサービスを守る。