生成AIセキュリティガイドライン Ver. 1.0（ひな形）

(コピーしてお使いください)

第1章 総則

第1条（目的）

本ガイドラインは、当社における生成AI（Generative Artificial Intelligence）の利用に関して、遵守すべき事項を定めるものである。業務効率化や新たな価値創造といった生成AIの恩恵を最大化しつつ、関連するセキュリティリスクを適切に管理し、以下の達成を目的とする。

1. 生成AIの安全かつ責任ある利用の促進
2. 法令、各種規格、および社内規程との整合性の確保
3. ステークホルダー（顧客、取引先、社会）に対する説明責任の遂行

第2条（適用範囲）

本ガイドラインは、当社の役員、正社員、契約社員、派遣社員、インターン、アルバイト、外部講師など、当社の情報資産に一時的または恒常的にアクセスするすべての者に適用する。

第3条（管理体制と相談窓口）

1. 統括責任者: 情報セキュリティ担当役員（CISO）は、本ガイドラインの遵守に関する全社的な責任を負う。
2. 推進部門: 情報システム部門は、本ガイドラインの運用、技術的対策の実施、および定期的な見直しを行う。
3. **AI**倫理委員会（または相談窓口）: 本ガイドラインの解釈や利用判断に迷う場合の相談窓口としてAI倫理委員会を設置する。
4. 規制変動モニタリング: 法務部門は、国内外のAI関連法規制の動向を半年ごとにモニタリングし、その結果をAI倫理委員会および推進部門へ報告する。

第2章 用語の定義

第4条（用語の定義）

本ガイドラインで用いる主な用語の意味は、次の通りとする。

1. 生成**AI**: テキスト、画像、音声、コード等を自律的に生成する能力を持つAI技術の総称。
2. モデル: 特定のタスクを実行するために学習されたAIの計算モデル。オープンソースモデルと商用APIモデルを含む。

3. プロンプト: 生成AIに指示や情報を与えるための入力テキストまたはデータ。

4. 機密情報: 個人情報、顧客情報、技術情報、財務情報など、漏洩した場合に当社または関係者に損害を与える可能性のある非公開情報全般。

5. ハルシネーション（幻覚）: 生成AIが、事実に基づかない、または文脈と無関係なもっともらしい情報を生成する現象。

6. プロンプトインジェクション: 攻撃者が巧妙なプロンプトを注入することで、AIの意図しない動作を引き起こさせ、機密情報を窃取したりシステムを乗っ取ったりする攻撃。

7. **RAG (Retrieval-Augmented Generation)**: 外部の信頼できる情報源を検索し、その結果を基に回答を生成することで、ハルシネーションを抑制し、回答の精度を向上させる技術。

第3章 倫理原則と責任ある**AI**の利用

第5条（基本倫理）

生成AIの利用にあたっては、人間の尊厳、プライバシー、公平性を尊重し、社会に貢献する目的で利用する。

第6条（禁止ユースケース）

理由の如何を問わず、以下の目的での生成AIの利用を固く禁止する。

- 違法行為の助長（マルウェアや爆発物の作成依頼など）
- 人権侵害、差別、ヘイトスピーチの生成・拡散
- 児童ポルノや性的搾取コンテンツの生成
- 偽情報（フェイクニュース）の意図的な作成・拡散
- 個人や組織への名誉毀損、嫌がらせ
- 著作権、商標権等の知的財産権の侵害

第7条（高リスク**AI**における人間による最終判断）

従業員の採用評価、人事考課、融資審査など、個人の権利や機会に重大な影響を及ぼす可能性がある「高リスクAI」に該当する用途で生成AIを利用する場合、**AI**による自動的な最終決定を禁止する。必ず人間によるレビューと最終判断を介在させなければならない。

第4章 利用サービスの管理

第8条（利用サービスの原則）

1. 生成AIの業務利用は、第6章に基づき会社が承認した「利用許諾サービスリスト」に掲載されたものに限定する。

2. 私物端末からのアクセス（BYOD）は、別途定めるBYOD利用規程に準拠する。なお、許可端末の具体的な技術要件（**MDM**登録、**VPN**接続等）は、別紙「**BYOD**端末チェックリスト」を参照のこと。

3. 未承認サービスの利用は原則禁止とする。

第9条（インシデント対応）

1. 報告義務と初動タイムライン: 生成AI利用に起因する情報漏洩やセキュリティインシデント（またはその懸念）を発見した者は、発見から30分以内に所属長および情報システム部門へ報告する。

2. 対応体制: 報告を受けた情報システム部門は、24時間以内に影響範囲の特定と封じ込め等の一次対応を完了させ、インシデント対応計画書に基づき行動する。

3. 責任分界点: 各部門の役割と責任の分界点は、別途定める「インシデント対応責任分界点図」に従う。

第**5**章 情報（データ）の取り扱い基準

第10条（情報の分類と入力制限）

生成AIに入力する情報は、その機密度に応じて分類し、取り扱いを厳格に管理する。

【テーブル形式】

分類	色	説明	入力可否	具体例
レベル**3**：機密情報	赤	漏洩が重大な損害を与える情報	入力絶対禁止	・個人情報、顧客の非公開情報 ・未公開の決算情報、M&A情報 ・**認証情報や秘密鍵を含むソースコード** ・アルゴリズム的機密（社外秘モデル等）
レベル**2**：社内情報	黄	社外非公開情報	要注意 （匿名化・一般化し、所属長の許可を得た場合のみ可）	・個人名等を削除した議事録 ・顧客IDを含むが個人特定はできない統計ログ

分類	色	説明	入力可否	具体例
				・**認証情報等を含まない社内向けツールや公開OSSのソースコードの一部**
レベル**1**：公開情報	青	公開済み、または一般的な情報	入力**OK** （※注）	・プレスリリース済みの情報 ・公開済みの仕様書 ・一般的な知識に関する質問
（※注）再識別リスクへの注意：公開情報同士の組み合わせで個人が推測できるリスクに常に留意すること。				

【箇条書きリスト形式（メール等での展開用）】

- レベル**3:** 赤（入力絶対禁止）
    - 内容: 漏洩が重大な損害を与える情報
    - 具体例: 個人情報、顧客の非公開情報、未公開の決算情報、認証情報や秘密鍵を含むソースコード、アルゴリズム的機密など。
- レベル**2:** 黄（要注意・匿名化等が必要）
    - 内容: 社外非公開情報
    - 具体例: 個人名等を削除した議事録、個人を特定できない統計ログ、認証情報を含まないソースコードの一部など。
- レベル**1:** 青（入力**OK**）
    - 内容: 公開済み、または一般的な情報
    - 具体例: プレスリリース、公開仕様書、一般的な知識の質問など。（※再識別リスクに注意）

第6章 モデルのライフサイクル管理

第11条（サービスの選定と評価）

（旧第10条から項目を抜粋・変更なし）

第12条（利用許諾リストと撤退手順）

（旧第11条から変更なし）

第13条（モデルの管理と廃棄）

1. 導入・バージョン管理: 利用するモデルの名称、バージョン、ライセンス、導入日を台帳で管理する。

2. 更新・テスト: モデルの更新やファインチューニングを行う際は、変更管理プロセスに従う。その際、社内評価データセットを用いて性能を評価し、事前に定めた基準（例: **BLEU** スコア ≧ **0.5**、**BERTScore** ≧ **0.9**）を満たすことを確認する。

3. 廃棄: 利用を終了したモデルや関連データは、安全かつ完全に消去する手順を定め、実行記録を保管する。その際、クラウドストレージ上のスナップショットやバックアップ領域からの削除漏れがないよう、二重に確認を行う。

第7章 オープンソースモデルの利用管理

（旧第13条、内容変更なし）

第8章 技術的セキュリティ対策

第14条（アクセス管理）

（変更なし）

第15条（監視とログの管理）

1. ログ管理: プロンプトと生成物のログは最低2年間、暗号化して保管する。

2. 暗号鍵の管理: ログの暗号化に用いる鍵は、年**1**回以上の頻度でローテーションする。鍵操作には管理者**2**名以上の多重署名を必須とする。

3. 改ざん検知: 生成物にはハッシュ値を付与し、保管・転送時の改ざんを検知できる仕組みを導入することを推奨する。

4. 自動アラート: 不審なアクティビティを検知するため、自動アラート閾値を設定する。

5. **RAG**構成: RAGを採用する場合は、別途定める「RAG構成設計指針」（参照先: [社内ドキュメントリンク]）に従うこと。

第9章 出力物の利用管理と品質保証

第16条（出力物の二次利用）

1. 権利帰属の確認: （変更なし）

2. 事実検証（ファクトチェック）: 出力物を社外への公開や重要な意思決定に用いる際は、必ず人間による事実検証フローを経なければならない。本レビューは原則として**2**営業日以内に完了させることを目標とする。

第17条（ハルシネーション対策）

出力内容に疑義がある場合は、ハルシネーションを疑い、複数の情報源で裏付けを取る。その際、社内標準ツール（例**:** 社内ナレッジベースのセマンティック検索、契約済みの外部論文**API**等）を活用することを推奨する。

第10章 利用者の教育と訓練

第18条（定期研修とオンボーディング）

1. 全従業員に対し、本ガイドラインに関する研修を年1回以上実施し、新入社員のオンボーディングプログラムにも本研修を必須として組み込む。

2. 研修には、ハンズオン形式を取り入れ、実効性を高める。

3. 外部講師やベンダーが研修に参加する場合は、事前に法務部門提供の**NDA**（秘密保持契約）雛形を用いて契約を締結する。

第19条（継続的な啓発）

情報システム部門は、社内ポータル等に以下のコンテンツライブラリを整備し、継続的に啓発を行う。

- 禁止・要注意ワードリスト

- 安全なプロンプトの設計テンプレート

- 業界・部門別のFAQ

- 最新の脅威情報やインシデント事例

第11章 監査と脆弱性評価

第20条（内部監査と改善活動）

1. ログ監査: 情報システム部門は、利用ログの監査を月次で実施する。

2. ヒヤリハット共有: 四半期ごとに、各部門からヒヤリハット事例を収集・分析し、全社に共有する改善サイクルを運用する。

第21条（脆弱性評価と演習）

1. 第三者レビュー: （変更なし）

2. レッドチーム演習: 年1回、専門チームによるレッドチーム演習を実施し、システムの脆弱性を評価・可視化する。

3. 演習後のフォローアップ: 演習で発見された脆弱性や課題については、主管部門が改善計画書（対策内容、責任者、完了期限を明記）を作成し、演習後**1**ヶ月以内に**AI**倫理委員会へ提出する。

第**12**章 法令・規格の遵守

第22条（準拠性チェックリスト）

当社の事業に関連する以下の法令・規格等への準拠状況を定期的に確認し、対応策を講じる。（各社で具体的な内容を追記・修正する）

- [ ] 個人情報保護法 **/ GDPR**: 個人情報の取り扱い、越境移転に関する要件

- [ ] 著作権法: 入力データおよび生成物の著作権に関する要件

- [ ] 各種ソフトウェアライセンス: オープンソースモデル利用時のライセンス遵守

- [ ] **NIST AI Risk Management Framework 1.0**: AIリスク管理のフレームワークへの準拠

- [ ] **ISO/IEC 42001 (AI**マネジメントシステム**)**: AIガバナンス体制の構築

☐ その他、業界固有のガイドライン（例：金融、医療）

第**13**章 罰則と報告の促進

第23条（懲戒処分）

1. 本ガイドラインへの違反が確認された場合、就業規則に基づき、段階的な処分を科す。

2. 初回の軽微な違反には、原則として懲戒処分ではなく教育的指導を行う。指導内容には、再研修の受講、理解度テストの再合格、再発防止策を記載した計画書の提出などを含む。

3. 再発する場合や重大な違反に対しては、就業規則に基づき、譴責、減給、出勤停止等の懲戒処分を検討する。

4. 故意または重大な過失による違反行為により会社に損害を与えた場合、損害賠償請求を行うことがある。

第**24**条（自己申告による減免）

インシデントの隠蔽を防ぎ、早期発見・対処を促進するため、意図しない情報漏洩や誤操作を速やかに自己申告した者に対しては、懲戒処分を減免または免除することがある。本措置は、事案の重大性や本人の協力姿勢を総合的に勘案し、AI倫理委員会が判断する。

附則

- このガイドラインは、YYYY年MM月DD日より施行する。

- 改訂時の通知フロー: 本ガイドラインの重要な改訂が行われた場合、法務部門および情報システム部門は、その内容を全従業員に対し、社内ポータルへの掲示および全社一斉メール（または**Slack**等のチャットツール）にて通知する。

終わりに── 次のステップと個別相談のご案内

1. 新しい挑戦を”小さなミス” で終わらせないために

- **Step** 0：本ひな形を参考に自社独自の”**AI**活用ガイドライン”を製作
- **Step** 1：作るだけで終わらず、”機密情報や個人情報を入力しない”、”著作物の著作権を確認する”等の重要事項をピックアップして社員に周知
- **Step** 2：定期的にセキュリティ教育を行い、社員のセキュリティ意識を醸成

まずは小さく初めて徐々に範囲を広げていきましょう！

2. 個別相談（無料）で伴走サポート

「自社の場合は何から始めれば良い？」
「セキュリティ教育や活用方法を社員にわかりやすく伝えたい…」
そんな時は、以下の日程共有リンクから**30**分のオンライン相談をご予約ください。

👉**https://app.spirinc.com/patterns/availability-sharing/CuveLAwUnplEYK7KBzUsv/confirm**

3. LINE公式アカウントで最新情報を受け取る

- **AI**活用の成功事例やテンプレートを配信
- 質問はチャットでいつでも受付
  👉 **https://lin.ee/xnL5k5G**

「まだ早いかも…」と感じたその時こそ、始めどき。
「今」の小さな行動が、これから訪れる「AI失業時代」に御社が生き残る力になります！