取組項目の全体像

# 国内外の政策・標準を参照し5つの取組領域、全36の取組項目を整理

| | 全社・組織としての対応を要する項目 | | | ユースケースに応じて内容が変化する項目 | 全社・組織としての対応を要する項目 |
|---|---|---|---|---|---|
| 区分<br>分類 | ルール・プロセスの明文化 (6) | 周知徹底・人材育成 (5) | 組織体制整備 (5) | 各リスク領域への対策 (10) | 透明性・アカウンタビリティ確保 (10) |
| ver1.0での取組項目 | 1 AIポリシー<br>2 AI利用ルール<br>3 AI管理ルール<br>11 ガバナンス全体の不断の見直し<br>12 リスクベースアプローチ<br>13 ハイリスクなAIへの対応 | 4 AIユースケースの把握・整理<br>5 社内ルールの浸透<br>34活用段階に応じた人材要件の定義<br>35 AIリスクについての周知徹底<br>36 AIリスク管理を担う人材の育成 | 6 経営層のオーナーシップと役割分担の明確化<br>7 AIガバナンス構築の戦略的な位置付け<br>8 関連領域の知見の集積・交流<br>9 社内における客観的な視点の確保<br>10 インシデント発生時の対応フローの確立 | 14 個人情報・個人データの適正な取得<br>15 個人情報・個人データの適正な利用<br>16 データ取得における知的財産権の尊重<br>17 AIの利用における知的財産権の尊重<br>18 AIの設計における人間との相互作用のあり方、多様性・包摂性の考慮<br>19 AIの精度低下や出力の誤りへの対策<br>20 AIのバイアス・倫理的に問題のある出力への対策<br>21 AIの悪用への技術的対策<br>22 AIに対する攻撃への技術的対策<br>23 AIの悪用や攻撃へのルール設計による対策 | 24 継続的なリスク管理<br>25 組織やルールに対する外部の視点からの検証<br>26 第三者による技術面のリスク検証<br>27 ステークホルダーとの対話<br>28 バリューチェーン全体でのガバナンス実現<br>29 透明性の確保<br>30 アカウンタビリティの確保<br>31 監査可能性の確保<br>32 データのトレーサビリティの確保<br>33 環境・持続可能性への影響の考慮 |

回答完了 回答完了

| 取組区分<br>分類 | No. | 項目名 | 開発 | 提供 | 利用 | 取組事項 | 評価 | 取組状況の概要 |
|---|---|---|---|---|---|---|---|---|
| ルール・プロセスの明文化 | 1 | **AIポリシー** | ○ | ○ | ○ | AIモデル・サービスがもたらしうる便益やリスク、自社のAI習熟度などを考慮しつつ、自社のAI活用における考え方の大枠を体外的に示すためのルール（以下、「AIポリシー」）を策定している | | |
| ルール・プロセスの明文化 | 2 | **AI利用ルール** | | | ○ | 生成AIサービス等のAIモデル・サービスを社員が利用する際に、各サービスの利用マニュアルとは別途、AIポリシー実現とリスクの低減のために組織として準拠すべきルールやガイダンス（以下、「AI利用ルール」）を策定している | | |
| ルール・プロセスの明文化 | 3 | **AI管理ルール** | ○ | ○ | | 自社の事業においてAIモデル・サービスを開発・運用する際に、AIポリシー実現とリスクの低減のために準拠すべきルールやガイダンス（以下、「AI管理ルール」）を策定している | | |
| 周知徹底・人材育成 | 4 | **AIユースケースの把握・整理** | ○ | ○ | ○ | 自社でAIリスク管理や情報システム管理にあたる部門において一括して、あるいは各事業部における分散管理等によってAIモデル・サービスのレジストリを整備し、自社内のAIモデル・サービスを網羅的に把握している | | |
| 周知徹底・人材育成 | 5 | **社内ルールの浸透** | ○ | ○ | ○ | 策定したAI利用ルールやAI管理ルールについて社員に周知徹底し、照会対応、担当者向けの説明会・相談会の実施等のルールの理解を促す活動も必要に応じて実施し、現場で遵守されるように徹底している | | |
| 組織体制整備 | 6 | **経営層のオーナーシップと役割分担の明確化** | ○ | ○ | ○ | AIガバナンスを担当する経営層がアサインされ、そのオーナーシップのもとで担当部局を決定してAIガバナンス施策を推進している | | |
| 組織体制整備 | 7 | **AIガバナンス構築の戦略的な位置付け** | ○ | ○ | ○ | AIガバナンスの構築を、AIの活用戦略と一体となる戦略的なアジェンダとして社内で取り上げ、部門の年次計画、経営への報告事項などにおいて明確に位置付けている | | |
| 組織体制整備 | 8 | **関連領域の知見の集積・交流** | ○ | ○ | ○ | AIガバナンスに関連する諸領域の知見を持つ部門・人（技術、セキュリティ、法令、コンプライアンス等）同士が協業したり、最新の制度やAIリスクや攻撃手法の動向、AIリスク管理のベストプラクティス等について知見交流を行う社内外の組織体や議論の場を確保している | | |
| 組織体制整備 | 9 | **社内における客観的な視点の確保** | ○ | ○ | ○ | AIモデル・サービスのリスク管理プロセスにおいて、その活用を推進する部門・人のみでなく、そのAIモデル・サービスの開発・運用に携わっておらず客観的な視点をもつ部門・人がダブルチェックを行う体制となっている（いわゆる金融分野での「3つの防衛線」の考え方における2線による検証・牽制機能が体制として確保されている） | | |

回答完了 回答完了

| 取組区分<br>分類 | No. | 項目名 | 開発 | 提供 | 利用 | 取組事項 | 評価 | 取組状況の概要 |
|---|---|---|---|---|---|---|---|---|
| 組織体制整備 | 10 | **インシデント発生時の対応フローの確立** | ○ | ○ | ○ | AIモデル・サービスのリスクが顕在化してしまった場合（インシデント）に備えて、対応する責任者や組織の特定、対応フローの確立、説明ポリシーの設定といった体制の整備を行なっている | | |
| ルール・プロセスの明文化 | 11 | **ガバナンス全体の不断の見直し** | ○ | ○ | ○ | AIガバナンスにおけるマネジメントシステムや各ユースケースのリスク事前評価、開発・運用にあたってのリスクの検証・軽減手法などについて、一定期間の経過や法令規制の改正等のイベントをトリガーとした見直しを行い、最新の技術・社会動向を反映している | | |
| ルール・プロセスの明文化 | 12 | **リスクベースアプローチ** | ○ | ○ | ○ | 自社のAIモデル・サービスについて、開発・運用の本格着手に先立ち網羅的にリスクの事前評価を行い、各モデル・サービスがどの程度のリスク・レベルを持ち、どのようなリスク管理を必要としているかを精査している | | |
| ルール・プロセスの明文化 | 13 | **ハイリスクなAIへの対応** | ○ | ○ | ○ | リスクの事前評価でAIの出力結果を用いた判断・業務遂行が人間の権利利益の侵害に繋がりうる等の要件によりハイリスクと判断されたAIモデル・サービスについては、適切な人の関与を確保する仕組み（Human in the Loop）や、一定の問題が出た際に自動での運用を止める仕組み（Human over the Loop）などの導入を検討している | | |
| 各リスク領域への対策 | 14 | **個人情報・個人データの適正な取得** | ○ | ○ | ○ | AIモデル・サービスの学習データ、入力データ、検索拡張生成等に用いるデータベース用のデータとして個人情報・個人データを取得する場合などに、同意取得などの適切な手続きを行う仕組みを構築している<br>＞本事項に取り組まない場合、法令（個人情報保護法）違反、適正な手続の欠如によるユーザからの苦情・訴訟等に発展するおそれがある | | |
| 各リスク領域への対策 | 15-1 | **個人情報・個人データの適正な利用（開発・提供者）** | ○ | ○ | | AIモデル・サービスによる個人情報・個人データの処理にあたり、個人情報保護法に違反する利用や目的と関連性のないデータの利用、プライバシー侵害にあたる利用が発生しないよう、適切な手続きを行う仕組みを構築している<br>＞本事項に取り組まない場合、法令（個人情報保護法）違反、AIモデル・サービスによる不適切な出力、適正な手続の欠如によるユーザからの苦情・訴訟等に発展するおそれがある | | |
| 各リスク領域への対策 | 15-2 | **個人情報・個人データの適正な利用（利用者）** | | | ○ | AIモデル・サービスによる個人情報・個人データの処理にあたり、個人情報保護法に違反する利用や目的と関連性のないデータの利用、プライバシー侵害にあたる利用が発生しないよう、適切な手続きを行う仕組みを構築している<br>＞本事項に取り組まない場合、法令（個人情報保護法）違反、AIモデル・サービスによる不適切な出力、適正な手続の欠如によるユーザからの苦情・訴訟等に発展するおそれがある | | |
| 各リスク領域への対策 | 16 | **データ取得における知的財産権の尊重** | ○ | ○ | ○ | AIモデル・サービスの学習データ、入力データ、検索拡張生成等に用いるデータベース用のデータを取得する際に、著作物や他者の登録商標等を適法に利用するための仕組みを構築している<br>＞本事項に取り組まない場合、法令（著作権法、意匠法、商標法等）に定められた権利の侵害に発展するおそれがある | | |

回答完了　　回答完了

| 取組区分<br>分類 | No. | 項目名 | 開発 | 提供 | 利用 | 取組事項 | 評価 | 取組状況の概要 |
|---|---|---|---|---|---|---|---|---|
| 各リスク領域への対策 | 17 | **AIの利用における知的財産権の尊重** | ○ | ○ | ○ | AIモデル・サービスの出力が著作権などの知的財産権侵害にあたるコンテンツを含む可能性を極力減らす仕組みや、含んでも外部への直接提供や公表を止められる仕組みを構築している<br>＞本事項に取り組まない場合、法令（著作権法、意匠法、商標法等）に定められた権利の侵害や、権利侵害に当たらなくとも利害関係者からの苦情・訴訟等に発展するおそれがある | | |
| 各リスク領域への対策 | 18 | **AIの設計における人間との相互作用のあり方、多様性・包摂性の考慮** | ○ | ○ | | AIモデル・サービスの設計段階から、AIモデル・サービスの人間との相互作用における負の影響や、多様性・包摂性について考慮に入れて開発・提供を行なっている<br>＞本事項に取り組まない場合、ユーザによるAIの不適切な擬人化、嫌悪、過度な依存のリスクが顕在化したり、ユニバーサルデザインの観点やアクセシビリティを欠くサービスになるおそれがある | | |
| 各リスク領域への対策 | 19-1 | **AIの精度低下や出力の誤りへの対策（開発・提供者）** | ○ | ○ | | AIモデル・サービスの開発・運用にあたり、その精度の低下や生成AIにおけるハルシネーションなどの出力の誤りが発生しうることを認識し、開発時・運用時それぞれにおいて、リスク検証と対策を行なっている<br>＞本事項に取り組まない場合、AI搭載製品の誤作動や誤りを含んだ手続の遂行の結果、個人の生命・身体や権利利益の侵害などにつながり、当該AIモデル・サービスの提供先から責任を追求される事態に発展するおそれがある | | |
| 各リスク領域への対策 | 19-2 | **AIの精度低下や出力の誤りへの対策（利用者）** | | | ○ | AIモデル・サービスにおいては、精度の低下や生成AIにおけるハルシネーションなどの出力の誤りが発生しうることを認識し、出力の正確性を確保するためのダブルチェックなどの対策を行なっている<br>＞本事項に取り組まない場合、AI搭載製品の誤作動や誤りを含んだ手続の遂行の結果、個人の生命・身体や権利利益の侵害などにつながり、業法への違反や、影響を受けた顧客等からの苦情・訴訟等に発展するおそれがある | | |
| 各リスク領域への対策 | 20-1 | **AIのバイアス・倫理的に問題のある出力への対策（開発・提供者）** | ○ | ○ | | AIモデル・サービスの開発・運用にあたり、出力のバイアスや、倫理的に問題のある出力を抑制するため、開発時・運用時それぞれにおいて、リスク検証と対策を行なっている<br>＞本事項に取り組まない場合、学習データやアルゴリズムに起因する出力のバイアスや、差別語・ハラスメント等の倫理的に問題のある出力が発生し、当該AIモデル・サービスの提供先とのトラブルに発展するおそれがある | | |
| 各リスク領域への対策 | 20-2 | **AIのバイアス・倫理的に問題のある出力への対策（利用者）** | | | ○ | AIモデル・サービスの利用に当たって、出力のバイアスを排除するためのフィルタリングツールの導入やダブルチェックなどの対策を行なっている<br>＞本事項に取り組まない場合、学習データやアルゴリズムに起因する出力のバイアスや、差別語・ハラスメント等の倫理的に問題のある出力が発生し、誤った判断や問題のある内容の顧客への提示を通じたトラブルに発展するおそれがある | | |
| 各リスク領域への対策 | 21 | **AIの悪用への技術的対策** | ○ | ○ | ○ | AIモデル・サービスの開発・運用にあたり、悪意ある主体による問題あるコンテンツの出力を狙った入力や出力結果の悪用が発生しうることを認識し、開発時・運用時それぞれにおいて、リスク検証と対策を行なっている<br>＞本事項に取り組まない場合、当該AIモデル・サービスを悪用した偽情報の生成、個人の肖像や声を悪用したディープフェイクの作成、サイバー攻撃やテロなどの犯罪・CBRN（Chemical, Biological, Radiological, Nuclear）に利用されうる情報の生成などが容易になり、社会への悪影響につながるおそれがある | | |

回答完了 回答完了

| 取組区分分類 | No. | 項目名 | 開発 | 提供 | 利用 | 取組事項 | 評価 | 取組状況の概要 |
|---|---|---|---|---|---|---|---|---|
| 各リスク領域への対策 | 22 | **AIに対する攻撃への技術的対策** | ○ | ○ | ○ | AIモデル・サービスの開発・運用にあたり、開発時・運用時それぞれにおいて、データポイズニングや生成AIにおけるプロンプト・インジェクションといったAIシステムへの攻撃を念頭に、リスク検証と保護機能の導入や適切なアクセス権制御等の対策を行なっている<br>＞本事項に取り組まない場合、AIモデル・サービスの脆弱性（特にAIシステムに特徴的な脆弱性）を突く攻撃によって情報漏洩、モデル窃取、システム障害といった被害が生じるおそれがある | | |
| 各リスク領域への対策 | 23 | **AIの悪用や攻撃へのルール設計による対策** | ○ | ○ | | AIモデル・サービス調達の契約やサービスの利用規約等において、適切な用途の制限などを規定し、悪意ある主体による利用を防いでいる<br>＞本事項に取り組まない場合、悪用や攻撃による被害が生じた際に、AIモデル・サービスの開発・提供、業務利用者においての対策が不十分であったとして責任が追及されるおそれがある | | |
| 透明性・アカウンタビリティ確保 | 24-1 | **継続的なリスク管理（開発・提供者）** | ○ | ○ | | AIモデル・サービスの開発・運用にあたり存在するリスクは、データや技術環境、攻撃手法などの変化により変動していくことを踏まえ、開発時だけでなく運用中も不断にリスク検証と対策をアップデートしている<br>＞本事項に取り組まない場合、提供中のAIモデル・サービスに最新の保護策が適用されていないことにより、多数の提供先において同じ攻撃による問題が発生するなど、被害や提供先とのトラブルの規模が拡大するおそれがある | | |
| 透明性・アカウンタビリティ確保 | 24-2 | **継続的なリスク管理（利用者）** | | | ○ | AIモデル・サービスの利用にあたり存在するリスクは、データや技術環境、攻撃手法などの変化により変動していくことを踏まえ、開発時だけでなく運用中も不断にリスク検証と対策をアップデートしている<br>＞本事項に取り組まない場合、学習や検索拡張生成に用いる自社データによる悪影響の軽減等、AI利用者が主体となるべきリスク対策が疎かになり、運用中のAIモデル・サービス提供の質が低下したり問題が発生したりするおそれがある | | |
| 透明性・アカウンタビリティ確保 | 25 | **組織やルールに対する外部の視点からの検証** | ○ | ○ | ○ | 外部有識者や認証・監査主体のチェックなどを通じて、自社のAIマネジメントシステムが適切に機能することを確認している | | |
| 透明性・アカウンタビリティ確保 | 26 | **第三者による技術面のリスク検証** | ○ | ○ | | ハイリスクなユースケースのAIモデル・サービスを中心にするなどリスクレベルに応じて、その開発・運用に携わっていない社外の専門家・ツール提供者等も必要に応じて含んだ第三者によるレッド・チーミング等のリスク検証を実施している | | |
| 透明性・アカウンタビリティ確保 | 27 | **ステークホルダーとの対話** | ○ | ○ | | バリューチェーン上に位置するプレイヤー（AIモデル・サービスの開発者・提供者・利用者や、データの提供者等）に加え、AIモデル・サービスの利用者やサービスから影響を受ける者など、関連するステークホルダーとコミュニケーションするチャネルを作り、そうした意見をもとにAIモデル・サービスやAIガバナンスのあり方を改善している | | |
| 透明性・アカウンタビリティ確保 | 28 | **バリューチェーン全体でのガバナンス実現** | ○ | ○ | ○ | バリューチェーン上に位置するプレイヤー（AIモデル・サービスの開発者・提供者・利用者や、データの提供者等）同士、契約や自主的約束等により責任分界や他のプレイヤーに求める対応の明確化に努め、プライバシーや営業秘密を尊重しつつ必要かつ合理的な範囲で情報共有を行うなど、バリューチェーン全体でのガバナンス実現を考慮に入れた取組を行なっている | | |

回答完了 回答完了

| 取組区分<br>分類 | No. | 項目名 | 開発 | 提供 | 利用 | 取組事項 | 評価 | 取組状況の概要 |
|---|---|---|---|---|---|---|---|---|
| 透明性・アカウンタビリティ確保 | 29 | **透明性の確保** | ○ | ○ | ○ | 自社の開発したAIモデル・サービスや外部AIサービスの利用について、AIを利用している事実やAIの能力・限界、適切・不適切な利用方法といった事項を利用者やステークホルダー向けに開示している | | |
| 透明性・アカウンタビリティ確保 | 30 | **アカウンタビリティの確保** | ○ | ○ | ○ | 自社の開発したAIモデル・サービスや外部AIサービスの利用について、自社の取組や法令への準拠状況の説明・開示、及びAI活用をめぐる責任体制の明確化を行い、常に対外的なアカウンタビリティ（主に説明責任、答責性を含意）を果たせる体制を整備している | | |
| 透明性・アカウンタビリティ確保 | 31 | **監査可能性の確保** | ○ | ○ | ○ | 自社の開発したAIモデル・サービスや外部AIサービスの利用について、利用目的の文書化、開発・運用や利用の実績・ログの保存、活用にあたって実施したリスク評価や技術的な検証の結果の保存を行い、インシデント発生時等に外部からAIガバナンスの実態を確認できる情報の整理と保存を可能にするプロセス・システム・体制を整備している | | |
| 透明性・アカウンタビリティ確保 | 32 | **データのトレーサビリティの確保** | ○ | ○ | | AIモデルの学習や関連して活用するデータ等について、その出所や加工の経緯等について実績・ログを保存し、必要に応じて確認できるプロセス・システム・体制を整備している | | |
| 透明性・アカウンタビリティ確保 | 33 | **環境・持続可能性への影響の考慮** | ○ | ○ | ○ | AIモデル・サービスの運用における計算資源の集中的な使用やそれに関する活動による電力の消費等、開発・提供・利用において、ライフサイクル全体で、地球環境への影響を検討している | | |
| 周知徹底・人材育成 | 34 | **活用段階に応じた人材要件の定義** | ○ | ○ | ○ | 開発、提供、利用、ガバナンス等、AIとの関与の仕方によって人材の要件を定義し、必要に応じて外部の標準等も参照して、各人材がスキル獲得を推進できる状態を作っている | | |
| 周知徹底・人材育成 | 35 | **AIリスクについての周知徹底** | ○ | ○ | ○ | AIモデル・サービスの開発・提供・利用においてさまざまなリスクが存在することや、それに対して実施すべき機密情報の扱いの留意等のリスク対策について、AIモデル・サービスを利用しうる職員全体に対して社内研修等のチャネルで周知し、必要に応じて試験等で習得状況を把握している | | |
| 周知徹底・人材育成 | 36 | **AIリスク管理を担う人材の育成** | ○ | ○ | ○ | 自社のAIガバナンスを担う人材を確保するため、AIモデル・サービスの開発・提供やリスク管理を担う部門において、人材要件に応じた知見を備えた人材の育成と、社内における適正な配置を実施している | | |