2024 年 5 月 10 日　全 9 頁

AI 事業者ガイドラインの公表

AI ガバナンスの構築とその継続的な改善が期待される

金融調査部 研究員　矢田　歌菜絵

［要約］

- 2024 年 4 月 19 日に総務省および経済産業省は「AI 事業者ガイドライン（第 1.0 版）」を公表した。これは、内閣府 AI 戦略会議で取りまとめられた「AI に関する暫定的な論点整理」を受けて、総務省および経済産業省が既存の AI に関する諸ガイドラインを統合およびアップデートし、ソフトローとして新たに策定したものである。

- このガイドラインの対象者は、AI を直接事業に活用する事業者で、AI 開発者・AI 提供者・AI 利用者に分けられている。このガイドラインでは、AI の開発や利用における基本的な理念や原則等を示すだけでなく、AI ガバナンスの統一的な指針を示すほか、AI 開発者・AI 提供者・AI 利用者の各々が取り組むべき内容やその方向性について、リスクベースアプローチで示されている。

- AI を活用する事業者においては、このガイドラインを参考にして AI に関連した具体的な取組みを進めることが期待されている。このガイドラインは一度策定したらそこで完結するものでなく、AI ガバナンスの継続的な改善に向け適宜更新されるものであることから、AI を活用する事業者においても、自社の AI ガバナンスを継続的に改善していくことが求められるだろう。

AI の開発や利用に関する統合的なガイドラインの公表

2024 年 4 月 19 日に総務省および経済産業省は「AI 事業者ガイドライン（第 1.0 版）」[1]（以下、本ガイドライン）を公表した。公表に先立ち政府は 2024 年 1 月 20 日から「AI 事業者ガイドライン案」について意見募集を実施した[2]。意見募集は 1 カ月間実施され、法人・団体と個人併せて 4,000 件弱の意見が提出された[3]。本ガイドラインは、内閣府 AI 戦略会議で取りまとめられた「AI に関する暫定的な論点整理」（2023 年 5 月 26 日）を受けて、総務省および経済産業

[1] 経済産業省「AI 事業者ガイドライン（第 1.0 版）」（2024 年 4 月 19 日）
[2] 経済産業省「AI 事業者ガイドライン案」（2024 年 1 月 19 日）
[3] 経済産業省　第 3 回 AI 事業者ガイドライン検討会、第 27 回 AI ネットワーク社会推進会議、第 23 回 AI ガバナンス検討会 合同会議　資料 1「『AI 事業者ガイドライン案』に対するご意見及びその考え方」（2024 年 3 月 14 日開催）

省が既存の AI に関する諸ガイドライン[4]を統合およびアップデートし、法的拘束力のないソフトローとして新たに策定したものである。

制定の背景

AI の発展が急速に進み、イノベーションの創出や社会課題解決にも AI の活用が進む中、AI がもたらす社会的リスクの多様化や増大が進んでいる。AI によるリスクのごく一例として、偏ったデータに基づく学習の結果が特定の属性の人々に対して不利なものを示すおそれや、AI が生成したコンテンツが他者の著作物に酷似して著作権を侵害するおそれ等が挙げられる。

AI による多様なリスクに対する諸外国の対応を見ると、例えば、EU では AI 規則案の制定が予定されており、米国では AI の安全性に係る大統領令[5]が署名される等、法的拘束力を持った規制が整備されている。一方、日本では、AI の活用やリスクに対して法的拘束力を持った規制ではなく、ソフトローでの手当てがなされた。それが、本ガイドラインである。

日本において法的拘束力を持った規制ではなくソフトローでの手当てがなされた背景として、本ガイドラインでは以下の 3 つが指摘されている。

- 少子高齢化に伴う労働力の低下等の社会課題の解決手段として、AI の活用が期待されていること
- 法律の整備・施行が AI の技術発展及びその社会実装のスピード・複雑さとの間でタイムラグが発生すること
- 細かな行為義務を規定するルールベースの規制を行うと、イノベーションを阻害する可能性があること

（出所）本ガイドライン p.2

これらを踏まえ、AI によるリスクの低減とイノベーションの促進を両立させ、AI の開発や利活用に関わる全ての事業者による自主的な取組みを促すために、ソフトローでのゴールベースアプローチ（ゴールをまず明確にした上で具体的な取組みを定めること）を取った本ガイドラインを作成したとしている。

本ガイドラインの位置付け

本ガイドラインでは、AI の開発や利用における基本理念や原則等を示すだけでなく、AI ガバナンスの統一的な指針を示すほか、AI 開発者、AI 提供者、および AI 利用者の各々が取り組むべき内容やその方向性についてリスクベースアプローチ（予めリスクを特定し、そのリスクに対して必要な対応策を設定し、目標達成の確度を高めること）でまとめられている。AI の開発

[4] 総務省主導「国際的な議論のための AI 開発ガイドライン案」（2017 年 7 月 28 日）、「AI 利活用ガイドライン～AI 利活用のためのプラクティカルリファレンス～」（2019 年 8 月 9 日）、経済産業省「AI 原則実践のためのガバナンス・ガイドライン Ver. 1.1」（2022 年 1 月 28 日）

[5] 概要については拙稿「バイデン米大統領、AI の安全性に係る大統領令に署名」（2023 年 11 月 30 日付大和総研レポート）を参照されたい。

や利活用に関わる事業者においては、本ガイドラインを参考の一つとして、安心安全な AI の活用のための自主的かつ具体的な取組みを推進することが期待される。なお、本ガイドラインは一度策定したらそれで完結するものでなく、AI ガバナンスの継続的な改善に向け多様なステークホルダー（株主や取引先、顧客等のあらゆる利害関係者）の意見を踏まえて適宜更新されるものである（Living Document）。

本稿では、本ガイドラインの指針および、AI の開発や利活用に関わる事業者による説明に関連した「アカウンタビリティ」、「AI ガバナンス」について紹介する。

定義

本ガイドラインの対象や用語の定義については、本ガイドライン pp. 2-7「はじめに」や pp. 8-9「第 1 部 AI とは」に記載されている。本稿では主要なものを紹介する。

対象者

本ガイドラインは、「様々な事業活動において AI の開発・提供・利用を担う全ての者（政府・自治体等の公的機関を含む）」[6]を対象としている。AI 事業者の具体例は、「別添 1.『第 1 部関連 AI 事業者のパターン』」（別添 1 p. 10）に挙げられている。なお、AI を直接事業に活用せずに AI システム・サービスの便益を享受する（または損失を被る）者（図表 1 灰色箇所）についてはその対象には含めていない。

図表 1 本ガイドラインの対象者

<table>
<tr><td rowspan="3">様々な事業活動で
AIを活用
（政府・自治体等の
公的機関を含む）</td><td rowspan="3">AIを直接事業に活用</td><td>AI開発者</td><td>AIモデル等の開発、データ収集、前処理、
AIモデルの学習・検証、AIシステムの構築等</td></tr>
<tr><td>AI提供者</td><td>AIシステム検証、他のシステムへの実装、
AIシステム・サービス提供、運用支援等</td></tr>
<tr><td>AI利用者</td><td>AIシステム・AIサービスを利用</td></tr>
<tr><td colspan="2">AIを直接事業に活用せずに
AIシステム・サービスの便益を
享受する/損失を被る</td><td rowspan="2">業務外
利用者</td><td rowspan="2"></td></tr>
<tr><td colspan="2">事業活動以外でAIを活用</td></tr>
</table>

（注）青色箇所が対象者。
（出所）本ガイドライン pp. 4-5 を基に大和総研作成

AI とは

そもそも AI とは、Artificial Intelligence（人工知能）を指すが、確立された定義があるわけではない[7]とされている。本ガイドラインで AI と言及する際は、「『AI システム（中略）』自体

[6] 本ガイドライン p. 4
[7] 統合イノベーション戦略推進会議決定「人間中心の AI 社会原則」（2019 年 3 月 29 日）p. 1 参照。

又は機械学習をするソフトウェア若しくはプログラム含む抽象的な概念」[8]としている。ここでの AI システムとは、「活用の過程を通じて様々なレベルの自律性をもって動作し学習する機能を有するソフトウェアを要素として含むシステムとする（機械、ロボット、クラウドシステム等）」[9]とされている。AI サービスについては、「AI システムを用いた役務を指す」[10]としており、AI 利用者への価値提供全般を指している。そのほかにも、生成 AI といった高度な AI システムも本ガイドラインの対象とされ、その定義が記載されている。

AI の開発や利活用に関わる事業者が取り組むべき事項

本ガイドライン p.10 以降の「第 2 部 AI により目指すべき社会及び各主体が取り組む事項」においては、AI の利活用により目指す社会についての基本理念を記載した後、各主体が取り組む原則やその指針について記載されている。本稿では、基本理念と各主体に共通の指針、およびその指針から主体の説明責任に関する内容について紹介する。

基本理念

本ガイドラインでは、「人間中心の AI 社会原則」（2019 年 3 月 29 日）において示された基本理念を踏襲し、以下の 3 つを基本理念として掲げている。

①人間の尊厳が尊重される社会（Dignity）

②多様な背景を持つ人々が多様な幸せを追求できる社会（Diversity and Inclusion）

③持続可能な社会（Sustainability）

本ガイドラインにおいてもこれらの考え方は、技術の発展にかかわらず今後も目指すべき理念であり、尊重されるべき旨が記載されている。

共通の指針

共通の指針では、AI を活用して基本理念を実現するために各主体（AI 開発者、AI 提供者、および AI 利用者）が取り組むべき事項と、これらの主体が社会と連携して取り組むことが期待される事項に分類されている。

共通の指針は「安全性」や「公平性」といった「人間中心の AI 社会原則」を踏襲した 7 つの指針に、「教育・リテラシー」、「公正競争確保」、「イノベーション」の 3 つを加えた 10 の指針が掲げられている。共通の指針と主な内容は図表 2 の通りである。

各主体は人権の尊重や法令の遵守、多様性や公平公正な社会を尊重するための対応や、AI の

[8] 本ガイドライン p.8
[9] 本ガイドライン p.8
[10] 本ガイドライン p.9

開発・利用・提供に伴うリスクの分析・対策を講じる必要がある。例えば、AI による学習や出力にあたっては知的財産基本法や著作権法等の関連法令の遵守が、個人情報の保護にあたっては個人情報保護法や関連法案等の遵守がそれぞれ必要になる。また、本ガイドラインや他の国際的な指針、諸外国の法令やガイドライン等についても留意し、それらに則った対応が求められることになる。

図表 2 共通の指針とその主な内容

指針		主な内容
各主体が取り組む事項	1) 人間中心	・少なくとも憲法が保障するまたは国際的に認められた人権を侵すことがないようにすべき ・AIが人々の能力を拡張し、多様な人々の多様な幸せ（well-being）の追求が可能となるように行動することが重要 ・自動化バイアス等の AI に過度に依存するリスクに対して必要な対策を講じる ・AIが生成した偽情報・誤情報・偏向情報が社会を不安定化・混乱させるリスクに対して必要な対策を講じる ・社会的弱者によるAIの活用を容易にするよう注意を払う
	2) 安全性	・AIシステム・サービスの出力の正確性を含め、要求に対して十分に動作している（信頼性） ・パフォーマンスレベルを維持し、無関係な事象に対して著しく誤った判断を発生させないようにする（堅牢性） ・必要に応じて客観的なモニタリングおよび対処も含めて人間がコントロールできる制御可能性を確保する ・学習等に用いるデータの透明性の確保や法的枠組みの遵守、AIモデルの更新等を合理的な範囲で適切に実施する
	3) 公平性	・特定の個人ないし集団への不当で有害な偏見および差別をなくすよう努めることが重要 ・AIの出力結果が公平性を欠くことがないよう、AIに単独で判断させるだけでなく、適切なタイミングで人間の判断を介在させる利用を検討する
	4) プライバシー保護	・個人情報保護法等の関連法令の遵守、各主体のプライバシーポリシーの策定・公表等により、ステークホルダーのプライバシーが尊重・保護されるよう、その重要性に応じた対応を取る
	5) セキュリティ確保	・機密性・完全性・可用性を維持し、常時、AIの安全安心な活用を確保するため、その時点での技術水準に照らして合理的な対策を講じる ・新たな手法による外部からの攻撃のリスクに対応するための留意事項を確認する
	6) 透明性	・AIシステム・サービスの検証可能性を確保しながら、必要かつ技術的に可能な範囲で、ステークホルダーに対し合理的な範囲で情報を提供することが重要 ・AIの学習プロセス、推論過程、判断根拠等のログを記録・保存する
	7) アカウンタビリティ	・「共通の指針」の対応状況等について、ステークホルダーに対して、リスクの程度を踏まえ、合理的な範囲でアカウンタビリティを果たすことが重要 ・データの出所、意思決定等について、技術的に可能かつ合理的な範囲で追跡・遡求が可能な状態を確保する ・アカウンタビリティを果たす責任者を設定する ・必要に応じてリスク管理、安全性確保のためのAIガバナンスに関するポリシー、プライバシーポリシー等の方針を策定・公表する ・関係情報等を文書化して一定期間保管し、必要となった際に、入手可能かつ参照可能な状態とする
社会と連携した取組	8) 教育・リテラシー	・AIに関わる者が十分なレベルのAIリテラシーを確保するために必要な措置を講じる ・生成AIの活用の拡大による人間とAIの作業の棲み分けを受け、新たな働き方ができるよう教育・リスキリング等を検討する
	9) 公正競争確保	・AIをめぐる公正な競争環境の維持に努める
	10) イノベーション	・産学連携やオープンイノベーション等を推進する ・自社と他社のAIシステム・サービスの相互接続性と相互運用性を確保する ・標準仕様がある場合は、それに準拠する ・適切な情報提供を行う

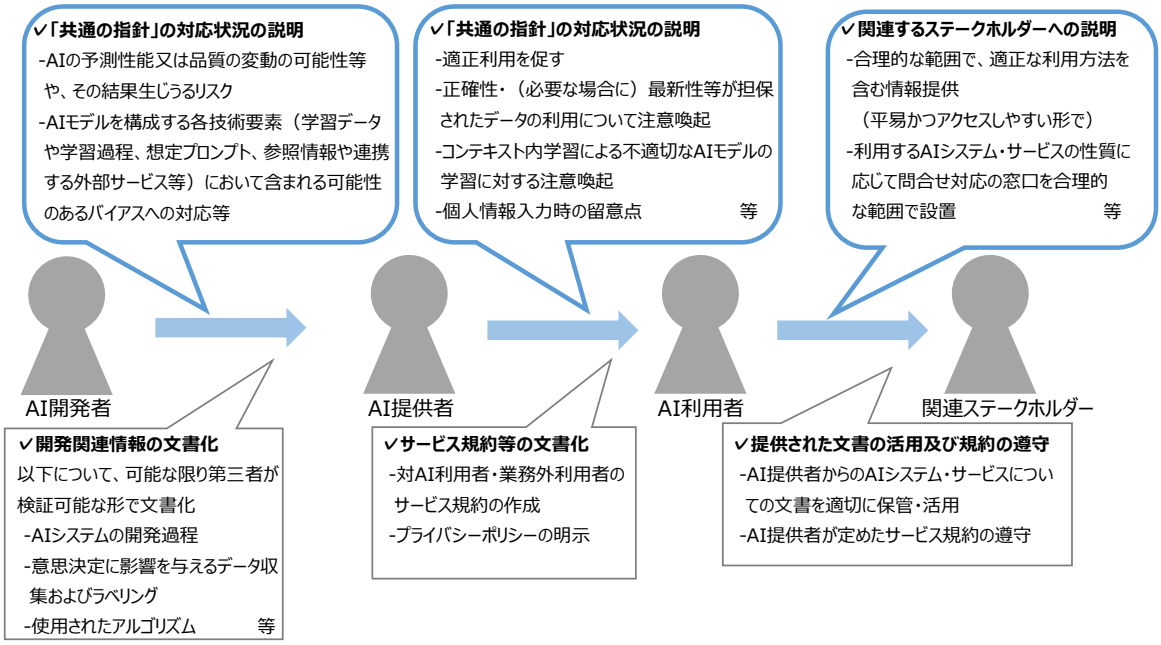
（出所）本ガイドラインを基に大和総研作成

アカウンタビリティ

前掲図表 2 の共通の指針において、各主体に取組みが期待されている事項の一つに「7)アカウンタビリティ」がある。一般に、アカウンタビリティは説明責任として用いられることが多いが、本ガイドラインでは、情報開示については「6)透明性」で対応し、アカウンタビリティの定義を「AI に関する事実上・法律上の責任を負うこと及びその責任を負うための前提条件の整備に関する概念とする」[11]としている。具体的には、ステークホルダーに対して、責任者の明示や、**共通の指針の対応状況の説明**、AI システム・サービスの開発・提供・利用における**意思決定等について**追跡等が可能な状態の確保することとしている。また、「7)アカウンタビリティ」では、プライバシーポリシーのみならず **AI ガバナンスに関するポリシーの策定および公表**も期待される取組み事項として挙げられている。

本ガイドラインおよび本ガイドライン別添にて例示されている「7)アカウンタビリティ」の具体的な取組みについて簡単にまとめると図表 3 の通りとなる。詳細は、本ガイドライン「第 3 部 AI 開発者に関する事項」、「第 4 部 AI 提供者に関する事項」、および「第 5 部 AI 利用者に関する事項」を参照されたい。

図表 3 各主体が取り組むべき主な事項（「7)アカウンタビリティ」の観点から）



（出所）本ガイドラインを基に大和総研作成

AI ガバナンス

前述の共通の指針「7)アカウンタビリティ」でもその方針の策定について言及されているが、共通の指針を実践し、AI を安全安心に活用するためには AI ガバナンスの構築が重要とされている。

[11] 本ガイドライン p. 18 脚注 21

本ガイドラインにおける AI ガバナンスとは、「AI の利活用によって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ、そこからもたらされる正のインパクト（便益）を最大化することを目的とする、ステークホルダーによる技術的、組織的、及び社会的システムの設計並びに運用」[12]と定義されている。AI のガバナンスは、その事業者内で完結するのではなく、AI の開発や提供、利用には様々な企業や機関が関わるため、多様なステークホルダーと連携させて実践していくことが特に重要とされている。

AI ガバナンス構築については、本ガイドラインにおいて、短期的な利益の追求の観点からではなく、中長期の課題および発展のための先行投資として捉えることが重要とされている。また、本ガイドライン「別添 2.『第 2 部 E.AI ガバナンスの構築』関連」にて記載されている留意点や実践のポイントでは、経営層が直接的に取り組むべき事項として例えば以下が挙げられ、一定の情報開示についてはコーポレートガバナンス・コードに係る情報に位置付けることの検討も含まれる。これらを考慮すると、AI ガバナンスもコーポレートガバナンスの一環として捉えられていることがうかがえる[13]。

1. 環境・リスク分析	✓ 経営層への報告・共有 ➢ 取締役会に対して責任を負う AI ガバナンスに関する社内組織の設置 ➢ AI ガバナンスに関する取組み状況を、取締役会において報告 等
2. ゴール設定	✓ 各主体の「AI ガバナンス・ゴール」を設定するかを検討 ➢ 設定する場合、合理的な範囲でそのゴールをステークホルダーに対して公開 ➢ 設定する場合、自社の取組み方針を含める 等
3. システムデザイン	✓ 現状の AI システム・サービス及び「AI ガバナンス・ゴール」からの乖離を特定・評価 ➢ リスクが認められる場合、その受容の合理性の有無を判定 ➢ 再考プロセスについて基本方針等を策定 ✓ 外部講師の招へいや教材の利用等で役職及び担当に適した AI リテラシー向上のための研修の実施 ✓ インシデントの予防及び早期対応 ✓ 責任の所在の明確化 等
4. 運用	✓ AI マネジメントシステムの運用状況について、関連するステークホルダーに説明可能な状態に ✓ AI ガバナンスに関する情報の透明性の確保 ➢ 開示する場合は、コーポレートガバナンス・コードの非財務情報に位置付けることも検討 等
5. 評価	✓ 継続的改善に向けた評価の重点ポイントを経営者自らの言葉で明示 ✓ AI マネジメントシステムの設計や運用から独立した関連する専門性を有する者の割り当て ✓ AI マネジメントシステムが適切に機能しているかモニタリング 等
6. 環境・リスクの再分析	✓ 適時の再評価等を行い、それに即した AI システムの改善や再構築等 ✓ AI ガバナンスの考え方を組織の文化として根付かせる 等

（出所）本ガイドライン別添を基に大和総研作成

[12] 本ガイドライン p. 9

[13] 例えば、小塚荘一郎「AI 原則の事業者による実施とコーポレートガバナンス」『情報通信政策研究』第 4 巻第 2 号（2021 年 3 月 25 日）pp. I-38-I-39 は、「（前略）AI 原則の実施やそのための AI ガバナンスの構築は、中長期的な株主の利益（『啓発された株主利益』）のために必要であり、会社経営者はそれを実行する義務を負うと言えそうである。」との見解を示している。

AI ガバナンスの構築にあたってのポイントや例は、本ガイドライン別添 2.「第 2 部 E.AI ガバナンスの構築」 関連（本ガイドライン別添 pp.18-70）に記載されており、AI ガバナンスの構築の際に参考とされるものである。また、本ガイドラインを基にした AI のガバナンスの検討にあたっては、「環境・リスク分析」「ゴール設定」「システムデザイン」「運用」「評価」というサイクルを、複数のステークホルダーで継続的に回していく「アジャイル・ガバナンス」[14]の実践が重要とされる（図表 4）。

アジャイル・ガバナンスとは、従来の中央集権的なガバナンス体制では急速かつ多様な変化に対応しきれないとの考えの下で提唱されたモデルで、次の 3 つの要素からなるものとされる[14]。

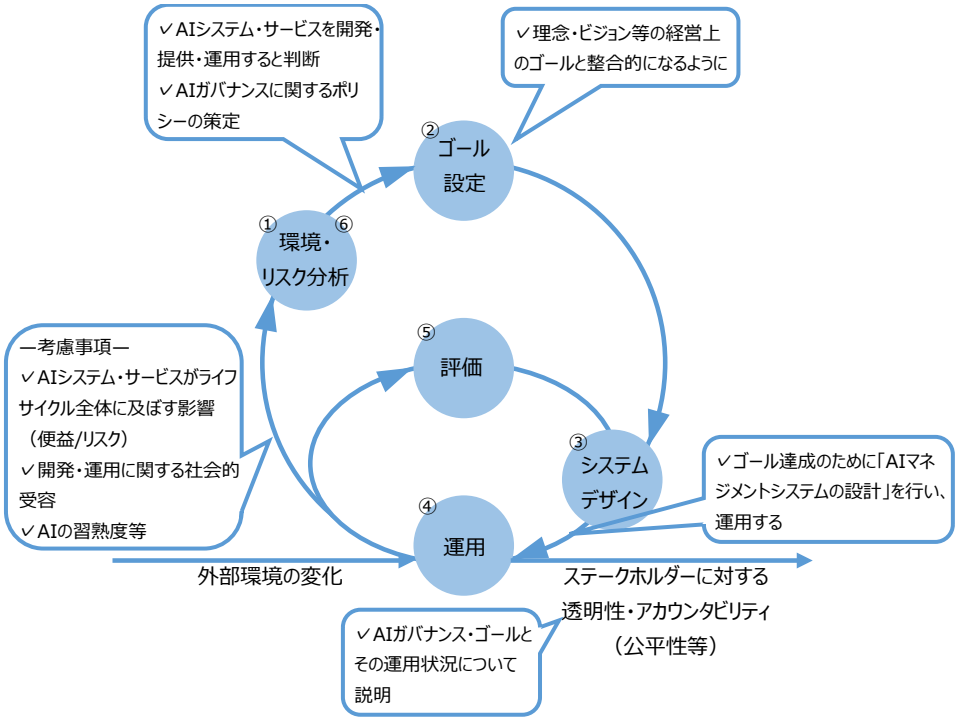
①主体：マルチステークホルダー（政府、企業、個人、コミュニティ等）

②手順：アジャイル（機動的かつ柔軟）

③構造：マルチレイヤー（各ステークホルダーに対するガバナンス）

アジャイル・ガバナンスにおいては、経営層レベルの外側のループと現場レベルの内側のループの 2 つを回していく必要がある。外側のループについては、一度ゴールを設定したら終わるのではなく、内側のループや外部環境の変化を受けて、現状のゴールやガバナンス体制が適切かどうか見直し続ける（図表 4 の「⑥環境・リスク分析」が該当）ことが特徴的である。

図表 4 アジャイル・ガバナンスの基本的なモデル（AI ガバナンスの観点から）



（出所）本ガイドラインおよび「Society5.0 における新たなガバナンスモデル検討会報告書（Ver.3）」を基に大和総研作成

また、AI の開発から利用までには多様な主体が関係することを踏まえ、バリューチェーンを

[14] 経済産業省　Society5.0 における新たなガバナンスモデル検討会「Society5.0 における新たなガバナンスモデル検討会報告書（Ver.3）」（2022 年 8 月 8 日）

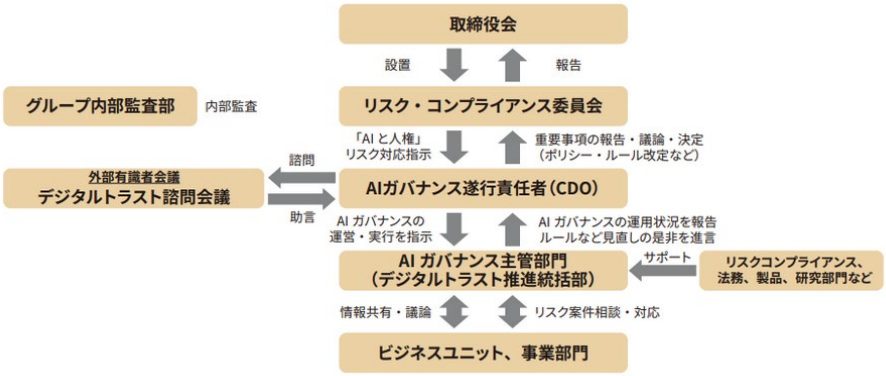
念頭においた検討が必要となる。そのバリューチェーンが複数国にまたがる場合も考えられ、そのような場合はデータの自由な越境移転のため、AI ガバナンスに係る国際社会の検討状況の把握や、それを踏まえた相互運用性についても確保することにも留意する必要があるとされている。

AI ガバナンスの取組み事例

本ガイドライン別添のコラムでは、AI ガバナンスに関する個社の事例を複数取り上げている。

例えば本ガイドライン「別添 2.『第 2 部 E. AI ガバナンスの構築」関連」（別添 pp. 63-64）においては、「コラム 5：NEC グループの AI ガバナンスに関する取組」を提示している（図表 5）。同グループでは、取締役会に加えて、AI ガバナンス遂行責任者（CDO：Chief Digital Officer）や、リスク・コンプライアンス委員会、外部有識者会議を置き、これらもコーポレートガバナンスに位置付けているという。

図表 5 NEC グループにおける AI ガバナンス体制



（出所）NEC グループ「AI と人権に関する方針」（最終閲覧日：2024 年 5 月 10 日）

おわりに

本ガイドラインは、AI ガバナンスの統一的な指針を示し、AI の開発・提供・利用にあたって必要な取組みについての基本的な考え方を示している。AI を活用する事業者においては、本ガイドラインを参考にして、AI に関連した具体的な取組みを進めることが期待されている。ガイドライン別添では主体ごとに取組みを進めるにあたってポイントとなる点が記載されている。事業者においては、自社の状況を把握した上で、これらのポイントも参考にしながら取組みを策定していくことが求められるだろう。

本ガイドラインは一度策定したらそれで完結するものでなく、AI ガバナンスの継続的な改善に向け適宜更新されるものである（Living Document）。ガイドラインの更新にあたっては、マルチステークホルダーで検討を重ねるとのことから、AI を活用する事業者における取組み状況も材料とされる。AI を活用する事業者においては、自社の AI ガバナンスを一度構築して満足するのでなく、改善に向けた取組みを継続することが求められるだろう。