AIガバナンスの概念と政府・企業・投資家による対応
－リスク調整後の投資パフォーマンスの向上に向けて－

江夏　あかね

▌要　約▐

1. 人工知能（AI）が急速に発展・普及する中で、リスクを管理・抑制しつつ最大限の便益を得ることを目的とするAIガバナンスの重要性が世界的に認識されてきた。
2. 各国・地域でスタンスは異なるものの、政府は法規制・ガイドライン等の対応を進め、企業自身も取締役会の監督体制や指針の整備、情報開示等の対応を進めている。さらに、近年は投資家も投資先企業のAI関連リスクに着目し、原則・ガイドラインの策定や情報開示等の対応を求めているほか、幅広いセクターの企業における対応状況に着目しつつある。
3. 金融資本市場の観点から、AIガバナンスを企業等にさらに浸透させていくことは、企業価値維持・向上、ひいてはリスク調整後の投資パフォーマンスの維持・向上にもつながり得るため、重要と言える。
4. その意味での今後の論点としては、（1）AIガバナンスの重要性に関する理解の促進、（2）情報・データの適切な開示、（3）第三者評価等の仕組みの整備、が挙げられる。特に、1点目について、米運用大手のフェデレーテッド・ハーミーズのように企業に対してAIに関するエンゲージメントを行う投資家が近年見られるようになっている。投資家が企業に対してエンゲージメントを実施するに当たっては、AIは、情報技術（IT）面のみの課題ではなく、企業価値に影響を及ぼし得る重要な経営課題であることを繰り返し伝え、企業の経営陣における意識向上を促すことが大切と言える。

野村資本市場研究所　関連論文等

・江夏あかね「欧州の証券監督当局が注視する証券市場におけるAIリスク－ESMAによる調査分析結果と今後のリスク対応の論点－」『野村資本市場クォータリー』第27巻第2号（2023年秋号）。
・橋口達「予測データ分析やAIの利活用に関する規制強化を図る米国SEC規則案－金融事業者と投資家間の利益相反への対応－」『野村資本市場クォータリー』第27巻第2号（2023年秋号）。

Ⅰ 世界的に活用が進むAIとそのガバナンスへの着目

人工知能（AI）[1]は近年、ビッグデータの発展、データストレージ容量の増加、計算機の処理能力向上等を背景に、世界で急速に発展・普及が進んでいる。とりわけ、米国のAI研究所であるOpenAIが開発した会話型AIサービス「ChatGPT」を始めとした生成AI[2]の進化を通じて、AIの活用可能性が大きく広がった。同時に、2018年前後からAI利用に伴う様々なリスクが徐々に顕在化し、AIに関するガバナンス（AIガバナンス）を確保する必要性が認識されるようになってきた[3]。AIガバナンスは、端的には、AI利用によるリスクを管理・抑制しつつ、最大限の便益を得ることを目的とするものである。適切なAIガバナンスは、企業を始めとした組織の価値を保全・向上する上で重要と言える。

近年は、AI利用によるリスクを管理・抑制すべく、国際機関、各国政府等が原則・ガイドラインを策定したり、法規制を強化しているほか、企業も自主的な取り組みを行うケースが増えている。足元の動きとして、国際連合のAIに関するハイレベル諮問機関は2024年9月、AI関連リスクとガバナンスのギャップに対処するための7つの提言を盛り込んだ最終報告書を公表した[4]。政府関連の動きとして、日本では、総務省・経済産業省が2024年4月に「AI事業者ガイドライン（第1.0版）」を公表した後、同年8月より内閣府の「AI制度研究会」にて制度のあり方について検討が始まった。米国では2023年10月、ジョー・バイデン大統領が「安全・安心・信頼できるAIの開発と利用に関する大統領令」を発令し、欧州連合（EU）では2024年5月、「AI規則」が成立し、同年8月に発効した。

本稿では、AIガバナンスの概念を整理するとともに、日米欧におけるAI関連規制等の現状及び企業や投資家による取り組み状況を概観した上で、金融資本市場の観点から今後の論点を考察する。

---

1 AIは、AIシステム（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）自体又は機械学習をするソフトウェア若しくはプログラムを含む抽象的な概念を指す（総務省・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 生成AIは、文章、画像、プログラム等を生成できるAIモデルに基づくAIの総称を指す（総務省・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）。

3 中崎尚『生成AI法務・ガバナンス―未来を形作る規範―』商事法務、2024年。

4 High-Level Advisory Body on Artificial Intelligence, “Governing AI for Humanity,” September 2024.

Ⅱ AIガバナンスの概念

本章では、（1）AI ガバナンスの定義とコーポレートガバナンスとの関係、（2）AI ガバナンスが求められる背景、を概観する。

１．AIガバナンスの定義とコーポレートガバナンスとの関係

AI ガバナンスに関して、世界共通の定義は存在しないとみられる。しかし、ガバナンス[5]の考え方をAI活用に当てはめると、AIのマイナス面を抑制するだけでなく、プラス面を活用することを目的として議論されることが多い傾向にある[6]。例えば、総務省・経済産業省による「AI 事業者ガイドライン（第 1.0 版）」では、AI ガバナンスを「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、社会的システムの設計並びに運用」と定義している。

AI ガバナンスとコーポレートガバナンス[7]との関係について、小塚（2021）は「AI 原則の実施も、AI の開発や利活用にかかわる企業においては、コーポレートガバナンスの課題として認識し、経営者（会社の形態に応じ、業務執行取締役ないし執行役）が必須の責務として取り組むべきもの」[8]との見方を明らかにしている。その背景として、（1）AI原則や AI 指針[9]の策定といった収益とは直結しないものの、事業者のリスクをコントロールする性格はサイバーセキュリティ対策に類似しており、経済産業省が策定した「グループ・ガバナンス・システムに関する実務指針」[10]では、サイバーセキュリティを内部統制システム上の重要項目として位置付けている、（2）経済産業省の検討会が公表した「我が国の AI ガバナンスの在り方 ver1.0」[11]と題した報告書では、AI 原則からガバナンスに進むことを提唱しており、関係する各事業者において、コーポレートガバナンスの中にAI原則の実施という課題を位置付ける必要性の指摘として理解される、との見解を挙げている。

---

5 ガバナンスは、組織の所有者が組織行動を制御するための仕組み。組織が目的達成に向けて適切に行動するように誘導し、その長期的な維持・存続・発展を可能にするために、取られる全ての統治・支配行動を指す（環境省「環境報告のための解説書－環境報告ガイドライン 2018 年版対応－」2019 年 4 月 12 日）。

6 前掲注 3 参照。

7 コーポレートガバナンスは、会社が、株主をはじめ顧客・従業員・地域社会等の立場を踏まえた上で、透明・構成かつ迅速・果断な意思決定を行うための仕組みを意味する（東京証券取引所「コーポレートガバナンス・コード ―会社の持続的な成長と中長期的な企業価値の向上のために―」2021 年 6 月 11 日）。

8 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、総務省情報通信政策研究所、2021 年。

9 小塚（2021）において、AI 原則は、AI の開発ないし利活用にあたっての考え方を示した文書のうち、各国の政府や民間団体などが策定するものを指している。AI 指針は、各事業者が、自社ないし自社グループにおける AI に対する考え方を示すものとして作成した文書を指している（前掲注 8 参照）。

10 経済産業省「グループ・ガバナンス・システムに関する実務指針（グループガイドライン）」2019 年 6 月 28 日。

11 経済産業省 AI 社会実装アーキテクチャー検討会「我が国の AI ガバナンスの在り方 ver. 1.0　AI 社会実装アーキテクチャー検討会　中間報告書」2021 年 1 月 15 日。

２．AI ガバナンスが求められる背景

AI は、1950 年代からの開発の歴史の中で進化が続き、企業活動や国民生活へ広く浸透してきた。特に、ディープラーニングの基盤技術により AI の性能が飛躍的に向上して誕生した生成 AI は、ユーザー側の調整やスキルなしに自然な言語で指示を出すだけで容易に活用できるといったこともあり、急速に普及が進んだ[12]。そして、後述のとおり、AI 活用に関する様々な便益やリスクが顕在化する中で、AI ガバナンスが求められていった。

企業にとっての AI 活用の主な便益としては、（1）運営コストの削減、（2）既存事業のイノベーションを加速させる新製品・サービスの創出、（3）組織変革、等が挙げられる[13]（図表 1 参照）。企業以外でも行政手続きの自動化、センサ及び画像情報等を用いた農場での作業支援システム、診療履歴等の利用による医療分野で活用が進んでいる。

図表 1　企業活動における AI による便益の例

分野	開発	マーケティング	販売	物流・流通	顧客対応	法務	ファイナンス	人事
従来から存在する便益の例（生成 AI でさらに向上）	コード検証、ドキュメント、作成の自動化	広告用メールの自動配信	受注後の対応メール等の自動発信	需要予測に基づく生産・在庫数	チャットボットによる自動対応	翻訳	財務諸表の自動作成	給与計算等の自動化
	類似のコード・データの抽出・検証	データに基づいたパーソナライゼーション	チャネル別、ニーズ別の売上予測	配送ルート最適化	過去の問い合わせ内容に基づいた FAQ 作成	法務文書のレビュー	過去実績に基づいた将来予測、不正検知	職務経歴書等に基づいた人材需要マッチング
生成 AI 特有の便益の例	学習データの生成、コーディングアシスタント、新製品のブレインストーミング	販売促進（マーケティング素材・キャッチコピー等）の自動作成	営業トークスクリプトの自動作成	物流条件交渉のアシスタント	対応内容の自動生成、要約	規定に基づいた契約書ドラフトの自動生成	文脈を踏まえた上での社内問い合わせ対応	文脈を踏まえた上での人事面接の対応

（出所）総務省・経済産業省「AI 事業者ガイドライン（第 1.0 版）別添（付属資料）」2024 年 4 月 19 日、より野村資本市場研究所作成

このように様々な便益が期待される AI だが、利用拡大や技術革新等により様々なリスクが顕在化している（図表 2 参照）。特に、生成 AI の普及をめぐっては、偽情報・誤情報の生成・発信等のリスクの多様化・増大が進む中、知的財産権の尊重を求める声等も聞かれるようになっている[14]。AI 利用に関するリスクは、法務、倫理、ビジネス上と多岐に渡っているほか、合法・違法の議論では割り切れないものも多く、レベル感にも幅がある。そのため、違法性リスクのように完全にゼロにすることを目指すのは現実的ではなく、リスクを管理すべく包括的なガバナンスが必要になると考えられている[15]。

[12] OpenAI による対話型 AI「ChatGPT」は、2022 年の発表後わずか 5 日で 100 万ユーザーを獲得し、さらに公開から 2 ヵ月後にはユーザー数が 1 億人を突破した（総務省「情報通信白書」2024 年 7 月）。
[13] 総務省・経済産業省「AI 事業者ガイドライン（第 1.0 版）別添（付属資料）」2024 年 4 月 19 日。
[14] 前掲注 13 参照。
[15] 前掲注 3 参照。

図表 2　AI によるリスクの例

項目	詳細
従来型 AI から存在するリスク	
バイアスのある結果及び差別的な結果の出力	・ 機械学習に使用した過去のデータ等が要因となり、バイアスのある結果や差別的な結果が出力されることがある
フィルターバブル及びエコーチェンバー現象	・ フィルターバブルとは、アルゴリズムがネット利用者個人の検索履歴やクリック履歴を分析し、学習することで、個々のユーザーにとって閲覧したい情報が優先的に表示され、利用者の観点に合わない情報からは隔離される情報環境のこと ・ エコーチェンバーとは、ソーシャルメディアネットワーク（SNS）等を利用する際、自分と似た興味関心を有するユーザーをフォローする結果、意見を SNS で発信すると自分と似た意見が返ってくること ・ AI 利用者及び業務外利用者が極端な考えを有する懸念がある
多様性の損失	・ 社会全体が同じモデルを、同じ用い方で使った場合、導かれる意見及び回答が大規模言語モデル（LLM）によって収束してしまい、多様性が失われる可能性がある
不適切な個人情報の取り扱い	・ 透明性を欠く個人情報の利用がある。例えば、人材採用に AI を用いるサービスにて、選考離脱及び内定辞退の可能性を AI により提供した際、学生等の求職者への説明が不明瞭であった他、一時的同意に基づいて第三者への情報提供が行われる規約になっていなかったこと等から個人情報保護法及び職業安定法にも違反することが判明し、サービスは廃止されることとなった ・ 個人情報の政治利用も問題視されている。例えば、SNS の業務外利用者に提供した「性格診断アプリ」及びプロフィール情報をもとに収集した個人情報を使用し、個々のパーソナリティを把握し、それに働きかけることで、依頼者に有利な投票行動をするようにターゲティング広告を打つ選挙支援活動が実施された
生命・身体・財産の侵害	・ AI が不適切な判断を下すことで、自動運転車が事故を引き起こし、生命及び財産に深刻な損害を与える可能性がある ・ インシデント発生時に優先順位付けを行うトリアージにおいては、AI が順位を決定する際に倫理的なバイアスを有することで、バイアスのある結果や差別的な結果が出力されることがある
データ汚染攻撃	・ AI 学習実施時には性能劣化及び誤分類につながるような学習データへの不正データ混入、サービス運用時にはアプリケーション自体を狙ったサイバー攻撃、AI の推論結果又は AI への指示であるプロンプト（AI との対話等においてユーザーが入力する指示や質問）を通じた攻撃等がリスクとして存在する ・ 例えば、あるチャットボットで、悪意のある集団による人種差別的な質問の組織的な学習により、ヘイトスピーチを繰り返し発言するようになってしまった
ブラックボックス化、判断に関する説明の要求	・ AI の判断のブラックボックス化に起因し、アルゴリズムの具体的な機能及び動作についての説明が難しいことがある
エネルギー使用量及び環境の負荷	・ AI の利用拡大により、計算リソースの需要が拡大し、結果として、データセンターが増え、そこでのエネルギー使用量の増加が懸念されている ・ ただし、例えばエネルギー管理に AI を導入することで、効率的な電力利用も可能になるなど、AI による環境への貢献の可能性がある
生成 AI で特に顕在化したリスク	
機密情報の流出	・ AI の利用において、個人情報及び機密情報がプロンプトとして入力され、その AI からの出力等を通じて流出してしまうリスクがある
悪用	・ AI の詐欺目的の利用が問題視されている。中でも、AI で合成された音声を利用した詐欺が急増している
ハルシネーション	・ 生成 AI が事実と異なることをもっともらしく回答する「ハルシネーション」に関しては AI 開発者・提供者への訴訟も起きている
偽情報、誤情報を鵜呑みにすること	・ 訴訟に関して、例えば、米国の弁護士が審理中の民事訴訟で資料作成に生成 AI を利用した結果、存在しない判例を引用してしまったことが問題となった ・ ディープフェイク（偽画像、及び偽動画を使った情報操作及び世論工作）に関する事例として、米国国防総省付近で爆発が起きたとする生成 AI で作られた偽画像が SNS 及びインターネットで瞬く間に拡散し、一部のメディアを装った偽アカウントも本情報を広げたことにより株価が一時大幅に下落するに至った
著作権との関係	・ アーティストの作品について、生成 AI に学習させ画像を生成した場合に、AI が学習した作品に似た画像が生成される場合があるとして、複数のアーティストが集団訴訟を起こした事例がある
資格等との関係	・ 生成 AI が法律または医療の相談に回答する場合、業法免許及び資格の侵害が生じ、法的問題が発生する可能性
バイアスの再生成	・ 生成 AI は既存の情報に基づいて回答を作るため、その答えを鵜呑みにする状況が続くと、既存の情報に含まれる偏見を増幅し、不公平及び差別的な出力が継続／拡大する可能性がある

（出所）総務省・経済産業省「AI 事業者ガイドライン（第 1.0 版）別添（付属資料）」2024 年 4 月 19 日、より野村資本市場研究所作成

Ⅲ 日米欧における AI 関連規制等の現状

AI リスク管理に関するルールの主な類型としては、（1）国際機関や各国政府が取りまとめた原理原則、（2）各国政府が定める中間的ルール、（3）企業が自主的に取り組む企業ルール、が挙げられる[16]（図表 3 参照）。本章では、このうち（1）と（2）の原理原則と中間的ルールを取り上げ、日本、米国、EU の状況を概観する（図表 4 参照）。これらの国・地域は、産業構成に加えて、文化や社会規範を背景とした AI に対する社会認識の違いに起因し、法規制に対するスタンスが異なる。しかしながら、後述のとおり、AI リスク管理を強化すべく取り組みを進めている面では共通していると言える（図表 5 参照）。

図表 3　AI リスクの積極的コントロールに関する規制等の枠組み

類型	詳細
原理原則	中立的な最終ゴール／最終ルール
	・ 人間中心、持続可能性、倫理といった AI の利用・開発に関するハイレベルな原則を定める ・ 国際機関が定めた原理原則である「AI に関する経済協力開発機構（OECD）原則」、「GPAI（Global Partnership on Artificial Intelligence）による実践的ガイダンス」や、各国政府が取りまとめた原理原則が存在
中間的ルール	分野横断的／個別分野ごとの中間的ルール
	・ 各国のガイドラインや、生体認証の利用や自動運転に関する既存規制の改正などの特定の利用や産業に関する個別規制 ・ 法的拘束力を伴うハードローと、ガイドラインに留まるソフトローが存在
企業ルール	各企業の自主的な取り組み（AI 活用／組織ガバナンス指針など）
	・ イノベーションを阻害せず企業による AI の設計や運用の達成手段をより具体化させる役割を期待されている ・ AI を活用したビジネスを展開するテック企業を中心にルールの策定・公表が始まっている

（出所）PwC「想定されるリスクと各国の法規制」2023 年 3 月 13 日、より野村資本市場研究所作成

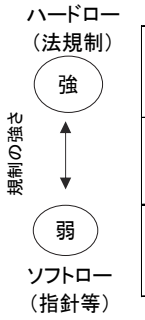
図表 4　AI に関連する主な諸原則等

時期	詳細
	日本
2017 年 7 月	総務省の AI ネットワーク社会推進会議、「国際的な議論のための AI 開発ガイドライン案」（AI 開発ガイドライン）を公表
2019 年 3 月	内閣府の統合イノベーション戦略推進会議、「人間中心の AI 社会原則」を決定
2019 年 8 月	総務省の AI ネットワーク社会推進会議、「AI 利活用ガイドライン」を公表
2022 年 1 月	経済産業省のAI原則の実践の在り方に関する検討会、「AI 原則実践のためのガバナンス・ガイドライン Ver. 1.1」を公表
2024 年 4 月	総務省と経済産業省、「AI 事業者ガイドライン（第 1.0 版）」公表
2024 年 8 月	内閣府、AI 制度研究会を開始
	諸外国
2019 年 4 月	EU の AI ハイレベル専門家グループ、「信頼性を備えた AI のための倫理ガイドライン」を公表
2019 年 5 月	OECD、「AI に関する理事会勧告」を採択
2021 年 11 月	国際連合教育科学文化機関（ユネスコ）、「AI の倫理に関する勧告」を採択
2022 年 10 月	米国ホワイトハウスの科学技術政策局（OSTP）、「AI 権利章典の青写真」を公表
2023 年 1 月	米国商務省の国立標準技術研究所（NIST）、「AI リスクマネジメントフレームワーク」（AI RMF）を公表
2023 年 10 月	米国のジョー・バイデン大統領、「AI の安全、安心、信頼できる開発と利用に関する大統領令」を公布
2023 年 12 月	G7、「広島 AI プロセス　包括的政策枠組み」を公表
2024 年 5 月	EU、AI 規則が加盟国に承認され、成立。同年 8 月に発効

（出所）総務省・経済産業省「AI 事業者ガイドライン（第 1.0 版）概要」2024 年 4 月 19 日、より野村資本市場研究所作成

[16] PwC「想定されるリスクと各国の法規制」2023 年 3 月 13 日。

図表 5　日米欧における AI 開発に関する規制等の動向



EU	- 人権重視でルールメーカーの地位を狙う -「AI規則」が成立（2024年5月）
米国	- 自主規制を主眼にしつつ法規制も -「安全・安心・信頼できるAIの開発と利用に関する大統領令」を発令（2023年10月）
日本	- 原則や指針に注力 -「AI事業者ガイドライン」公表（2024年4月） -「AI制度研究会」にて制度のあり方等について検討開始（2024年8月）

（出所）「海外の AI 関連企業、なぜ日本に進出するの？」『日本経済新聞』2024 年 6 月 8 日、より野村資本市場研究所作成

１．日本

日本については従来、「関係者による社会的リスクの低減を図るとともに、AI のイノベーション及び活用を促進していくために、関係者による自主的な取組を促し、非拘束的なソフトローによって目的達成に導く、ゴールベースの考え方で、ガイドライン[17]」の策定を中心に AI リスク管理を推進するスタンスであった。後述のとおり、法的拘束力の強いハードローを志向する EU とは異なるアプローチを採ってきたが、内閣府が 2024 年 8 月に開始した「AI 制度研究会」では、AI 利用によるリスクや影響も踏まえて制度の在り方について検討している。

本節では、（1）総務省・経済産業省が 2024 年 4 月に公表した「AI 事業者ガイドライン（第 1.0 版）」、（2）「AI 制度研究会」における検討状況、を中心に紹介する。

１）「AI 事業者ガイドライン（第 1.0 版）」

日本では、2016 年 4 月の G7 香川・高松情報通信大臣会合において AI 開発原則に向けた提案を行ったこと等も契機に、2010 年代後半から政府が各種取り組みを進めている。具体的には、総務省主導で「国際的な議論のための AI 開発ガイドライン案」（AI 開発ガイドライン）と「AI 利活用ガイドライン」及び経済産業省主導で「AI 原則実践のためのガバナンス・ガイドライン Ver. 1.1」を策定・公表してきた。また、内閣府が公表した「人間中心の AI 社会原則」を土台としつつ、上記のガイドラインを統合・見直すと共に、諸外国の動向や新技術の発展等を考慮して、総務省・経済産業省が 2024 年 4 月に「AI 事業者ガイドライン（第 1.0 版）」を公表した。

同ガイドラインでは、AI に携わる事業者が共通で遵守すべき原則を説明した上で、事業者を 3 つの主体（AI 開発者、AI 提供者、AI 利用者[18]）に分け、10 の項目から成

[17] 総務省・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。

[18] AI 開発者は、AI システムを開発する事業者（AI を研究開発する事業者を含む）。AI 提供者は、AI システムをアプリケーション、製品、既存のシステム、ビジネスプロセス等に組み込んだサービスとして AI 利用者、場合によっては業務外利用者に提供する事業者。AI 利用者は、事業活動において、AI システムまたは AI サービスを利用する事業者（前掲注 17 参照）。

る共通の指針（図表 6 参照）を示すとともに、3 主体それぞれに係る重要な項目も提示している。

AI ガバナンスの構築に関する留意点としては、（1）複数主体にまたがる論点について、バリューチェーン／リスクチェーンの観点で主体間の連携確保、（2）上記が複数国にまたがる場合、データの自由な越境移転の確保のための適切な AI ガバナンスの検討、（3）経営層のコミットメントによる、各組織の戦略や企業体制への落とし込み／文化としての浸透、が挙げられている。

また、同ガイドラインでは、サイバー空間とフィジカル空間を高度に融合させたシステム（CPS）を基盤とする社会は、複雑で変化が速いため、リスク統制が困難であり、社会の変化に応じて AI ガバナンスが目指すゴールが常に変化していくことを踏まえて、「アジャイル[19]・ガバナンス」の実践が重要と指摘している（図表 7 参照）。アジャイル・ガバナンスは「企業・法規制・インフラ・市場・社会規範といった様々なガバナンスシステムにおいて、『環境・リスク分析』『ゴール設定』『システムデザイン』『運用』『評価』といったサイクルを、マルチステークホルダーで継続的かつ高速に回転させていく」ものと説明されている。

図表 6　各主体（AI 開発者、AI 提供者、AI 利用者）に共通の指針

時期	詳細
各主体が取り組む事項	
人間中心	・ 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 ・ AI が生成した偽情報・誤情報・偏向情報が社会を不安定化・混乱させるリスクが高まっていることを認識した上で必要な対策を講じる ・ より多くの人々が AI の恩恵を享受できるよう社会的弱者による AI の活用を容易にするよう注意を払う
安全性	・ 適切なリスク分析を実施し、リスクへの対策を講じる ・ 主体のコントロールが及ぶ範囲で本来の利用目的を逸脱した提供・利用により危害が発生することを避ける ・ AI システム・サービスの特性及び用途を踏まえ、学習等に用いるデータの正確性等を検討するとともに、データの透明性の支援、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
公平性	・ 特定の個人ないし集団へのその人種、性別、国籍、年齢、政治的信念、宗教等の多様な背景を理由とした不当で有害な偏見及び差別をなくすよう努める ・ 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した上で、無意識や潜在的なバイアスに留意し、AI の開発・提供・利用を行う
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・公表により、社会的文脈及び人々の合理的な期待を踏まえ、ステークホルダーのプライバシーが尊重され、保護されるよう、その重要性に応じた対応を取る
セキュリティ確保	・ AI システム・サービスの機密性・完全性・可用性を維持し、常時、AI の安全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 ・ AI システム・サービスに対する外部からの攻撃は日々新たな手法が生まれており、これらのリスクに対応するための留意事項を確認する

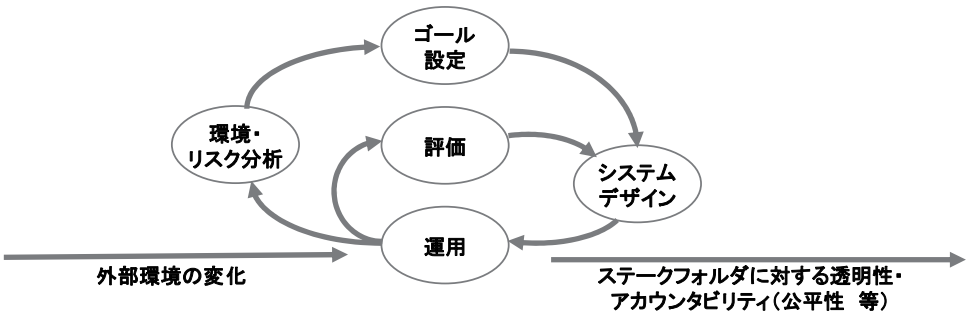
[19] アジャイル（agile）は、直訳すると、素早い、機敏なという意味である。システム開発やソフトウェア開発では、実装とテストを繰り返して開発していくという手法として用いられている（行政改革推進会議 アジャイル型政策形成・評価の在り方に関するワーキンググループ「アジャイル型政策形成・評価の在り方に関するワーキンググループ提言－行政の『無謬性神話』からの脱却に向けて－」2022 年 5 月 31 日）。

図表 6　各主体（AI 開発者、AI 提供者、AI 利用者）に共通の指針（続き）

時期	詳細
各主体が取り組む事項（続き）	
透明性	・ AI を活用する際の社会的文脈を踏まえ、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適切な情報を提供する（AI を利用しているという事実、活用している範囲、データ収集及びアノテーションの手法、AI システム・サービスの能力、限界、提供先における適切/不適切な利用方法、等）
アカウンタビリティ	・ トレーサビリティの確保や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情報の提供と説明を行う ・ 各主体の 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し、公表する ・ 関係する情報を文書化して一定期間保管し、必要なときに、必要なところで、入手可能かつ利用に適した形で参照可能な状態とする
社会と連携した取組が期待される事項	
教育・リテラシー	・ AI に関わる者が、その関わりにおいて十分なレベルの AI リテラシーを確保するために必要な措置を講じる ・ AI の複雑性、誤情報といった特性及び意図的な悪用の可能性もあることを勘案して、ステークホルダーに対しても教育を行うことが期待される
公正競争確保	・ AI を活用した新たなビジネス・サービスが創出され、持続的な経済成長の維持及び社会課題の解決策の提示がなされるよう、AI をめぐる公正な競争環境の維持に努めることが期待される
イノベーション	・ 国際化・多様化、産学官連携及びオープンイノベーションを推進する ・ 自らの AI システム・サービスと他の AI システム・サービスとの相互接続性及び相互運用性を確保する ・ 標準仕様がある場合には、それに準拠する

（注）　下線は原文による。
（出所）総務省・経済産業省「AI 事業者ガイドライン（第 1.0 版）概要」2024 年 4 月 19 日、より野村資本市場研究所作成

図表 7　アジャイル・ガバナンスにおける基本的なモデル



（出所）総務省・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）「AI 制度研究会」における検討状況

上述のガイドライン発出後、2024 年 6 月に閣議決定された「統合イノベーション戦略」に基づき、内閣府の「AI 戦略会議」の下に、AI 制度の在り方について検討することを目的として、「AI 制度研究会」が設置され[20]、2024 年 8 月から同研究会での検討が始まった。

2024 年 8 月 2 日に開催された AI 制度研究会第 1 回会合では、アジャイルでフレキシブルなガイドラインを最大限活用することを始めとした基本的な考え方等を基に、様々な有識者・専門家の意見を聴取し、議論を進め、2024 年秋に中間とりまとめを行うとの予定が示された。それと共に、事業主体を開発者、提供者・利用者に分けた上で、影響・リスクに応じた制度検討のイメージが提示されている（図表 8 参照）。

[20] 内閣府 AI 戦略会議「『AI 制度研究会』の設置について（案）」2024 年 7 月 19 日。

図表 8　AI 関係者を巡る制度検討のイメージ

事業主体	影響大・リスク大	影響小・リスク小
AI 開発者	① 確実なリスク対応 米国では大規模なモデルに報告義務 EU ではハイリスクな AI に様々な義務	② リスク対応 ルールを遵守していることの開示等
AI 提供者・利用者	③ 個別業規制等による基準遵守等 リスクの高い装置・機械類等の安全基準等	④ リスク対応 AI ガバナンスポリシーの策定・公表等
	⑤ 政府による適切な AI の調達・利用 リスクに関する情報の収集・公表	
プロバイダー	⑥ 不適切なコンテンツへの対応 オンラインプラットフォーマーによる違法情報への対応（EU のデジタルサービス法） テック企業による欺瞞的 AI 選挙コンテンツの削除等	

（出所）内閣府 科学技術・イノベーション推進事務局「AI 政策の現状と制度課題」2024 年 8 月 2 日、
より野村資本市場研究所作成

２．米国

ビッグテック企業を多く保有する米国は、自国の企業保護に力を入れ、政府による規制よりも民間での自主的な対応を優先し、企業の取り組みに任せつつ必要な規制をかけるというアプローチを採ってきた[21]。しかし、AI を取り巻くリスクの高まりを受けて、米国において AI 関連で初めての法的拘束力のある行政措置として位置付けられる[22]「安全・安心・信頼できる AI の開発と利用に関する大統領令」（AI 大統領令）が 2023 年 10 月に発令された。

本節では、連邦レベル[23]の動きに焦点を当て、（1）米国ホワイトハウスの科学技術政策局（OSTP）が 2022 年 10 月に公表した「AI 権利章典の青写真」、（2）米国商務省の国立標準技術研究所（NIST）が 2023 年 1 月に公表した「AI リスク管理フレームワーク」（AI RMF）、（3）AI 大統領令、について概観する。

１）「AI 権利章典の青写真」

OSTP は 2022 年 10 月、AI 開発等に当たり考慮すべき原則をまとめた「AI 権利章典の青写真」を公表した[24]。青写真では、AI を用いた自動化システムを設計、使用、配備する際に考慮すべき 5 つの原則が示された（図表 9 参照）。

21 総務省「令和 6 年版　情報通信白書」2024 年 7 月。
22 前掲注 3 参照。
23 州レベルでは、雇用場面における AI 技術の規制（アルゴリズムの規制、自動化された意思決定の規則）を中心に、基本的権利、何らかの検討体制を整備するための取り組みの観点からのルール整備が進んでいる（前掲注 3 参照）。
24 White House, “Blueprint for an AI Bill of Rights,” October 2022.

図表 9 「AI 権利章典の青写真」における 5 原則

原則	詳細
安全で効果的なシステム	システムは多様なコミュニティや利害関係者、専門家と協議の上、開発する。システムを配備する前に試験を行い、リスクを特定・軽減し、システムの監視を行う。これらの保護措置の結果として、システムの配備中止や削除もあり得る
アルゴリズムに基づく差別からの保護	システムが人種、性別、年齢などに基づいて不当な待遇をもたらすことがないよう、設計者、開発者、配備者はシステムを公平な方法で使用・設計するための継続的な措置を講じる
データ・プライバシー	個人の合理的な期待に合致し、厳密に必要なデータのみを収集する。システムの設計者、開発者、配備者は個人からの許可を取得し、データの収集、使用、アクセス、移転、削除に関する個人の決定を尊重する。個人の同意を求める際は、簡潔で、平易な言葉で理解できる内容にする。健康や仕事などに関わる機微なデータについては、より強い保護措置を講じる
通知と説明	システムの設計者、開発者、配備者は、システム全体の機能と自動化が果たす役割、そのようなシステムが使用されていることの通知、システムに責任を持つ個人・組織などを明確に説明する文書を広く一般に提供する。これらの情報は最新の状態に保ち、重要な使用例や主要機能の変更についてはシステムの影響を受ける人々に通知する
人間による代替、考慮、予備的措置	システムから影響を受ける個人が必要に応じてオプトアウトし、人間による代替手段を選ぶことができるようにする。システムの失敗やエラーが起きた場合などに、人間による考慮と予備的措置による救済を受けられるようにする

（出所）White House, “Blueprint for an AI Bill of Rights,” October 2022、「バイデン米政権、『AI 権利章典』を発表し AI 開発の 5 原則を示す」『ビジネス短信』日本貿易振興機構、2022 年 10 月 5 日、より野村資本市場研究所作成

同原則には、法的拘束力はないものの、OSTP は同原則が AI 製品の開発者や政策担当者、労働者等により米国社会で広く使用されることを期待するとともに、バイデン政権も連邦政府機関で同原則を推進する方針を示した[25]。

2）「AI リスク管理フレームワーク」（AI RMF）

NIST は 2023 年 1 月、AI 技術管理のためのガイダンスとして AI RMF を公表した[26]。同ガイダンスは、AI システムを設計、開発、導入、使用する者が自主的に参照可能となっており、AI 技術の効用を最大化しつつ、AI が個人や社会にもたらす悪影響を減らすためのリスク管理の手順がまとめられている[27]。また、上述の青写真と補完的関係になっている。

同ガイダンスは、（1）AI に関するリスクの考え方と信頼できる AI システムの特徴、（2）AI システムのリスクに対処するために取るべき対応策（「統治（組織におけるリスク管理文化の醸成等）」、「マッピング（リスクの特定）」、「測定（リスクの分析・評価等）、「管理（リスクの優先順位付けやリスクへの対応）」、に関して説明している。

3）「安全・安心・信頼できる AI の開発と利用に関する大統領令」（AI 大統領令）

AI RMF の公表以降、生成 AI の急速な浸透も背景として、民間側の自主的な取り組

25 「バイデン米政権、『AI 権利章典』を発表し AI 開発の 5 原則を示す」『ビジネス短信』日本貿易振興機構、2022 年 10 月 5 日。

26 National Institute of Standards and Technology, “NIST Risk Management Framework Aims to Improve Trustworthiness of Artificial Intelligence,” January 26, 2023.

27 「米商務省、AI のリスク管理のためのガイダンス発表」『ビジネス短信』日本貿易振興機構、2023 年 2 月 1 日。

みも進められてきた[28]。そのような中、バイデン大統領は 2023 年 10 月、AI 大統領令を発令した[29]。AI 大統領令は、AI がもたらす利益を享受するため、AI の無責任な使用による詐欺や差別、国家安全保障上のリスクを軽減することを目的としており、AI の安全性や連邦政府の AI 利用、国際的なリーダーシップ等、8 つの政策分野にわたって具体的な措置を記している[30]（図表 10 参照）。

対象とする AI の問題については、従来の倫理的観点から、安全保障問題に範囲を拡充し、対象となる事業者はビッグテック[31]に限らず、バイオテクノロジー企業等、国家の安全保障や経済に影響を及ぼす可能性のあるサービスや製品を取り扱う企業も含まれている[32]。

大統領令における内容面の主な特徴としては 2 点挙げられる[33]。1 点目について、従来型の AI リスクに関しては、前述の「AI 権利章典の青写真」の延長線にあるが、近い将来登場する可能性のある汎用 AI[34]（AGI）によってもたらされる恐れのある人類への脅威に対応すべく、デュアルユース（軍民共用）基盤モデルに対する評価・報告の義務付けや、化学・生物・放射線・核（CBRN）脅威におけるリスク軽減等、セキュリティを重視する傾向が強く見られている。

2 点目について、AI 大統領令の内容のほとんどは、連邦政府機関の行動を指示するものであるが、企業に対する強制力のあるルールの導入にもつながっている[35]。具体的には、商務省産業安全保障局（BIS）による、米国企業が外国の顧客にクラウドサービスを提供する際に、米国政府への報告等を義務付ける規則案（2024 年 1 月に官報で公示）が挙げられる[36]。なお、大統領令は、法律の裏付けがなければ将来の政権が自由に変更・破棄できるほか、大統領の取り組みを実行する上でも、議会による予算手当てが必要であること等を背景に、連邦議会へ提出される AI 関連法案の数は増加傾向にある[37]。

---

[28] AI 開発で先行する 7 社（グーグル、メタ、プラットフォーム、OpenAI 等）が AI の安全な開発のための自主的な取り組みを 2023 年 7 月に約束し、同年 9 月には新たな 8 社（IBM、アドビ、エヌビディア等）がそれに合意したことを米国政府が発表した。会社は、自主的なコミットメントとして、安全性、セキュリティ、信頼性の 3 つの観点から原則を掲げている（前掲注 21 参照）。

[29] White House, “Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,” October 30, 2023.

[30] 「AI 規制に大統領令で先手（米国）」『地域・分析レポート』日本貿易振興機構、2024 年 5 月 1 日。

[31] ビッグテックは、世界規模で支配的な影響力を持つ巨大 IT 企業群の通称。一般的には、米国のアルファベット、アップル、メタ、アマゾン、マイクロソフトの 5 社を指す（野村證券「ビッグ・テック｜証券用語解説集」）。

[32] 前掲注 21 参照。

[33] 前掲注 3 参照。

[34] 汎用 AI は、人間の知能の様々な側面を幅広くカバーし、様々な状況で人間の知能のように動作する汎用性の高いシステム（文部科学省「令和 6 年版 科学技術・イノベーション白書」2024 年 6 月）。

[35] 前掲注 30 参照。

[36] National Archives, “Taking Additional Steps To Address the National Emergency With Respect to Significant Malicious Cyber-Enabled Activities,” *Federal Register*, January 29, 2024、「米商務省、外国企業へのクラウドサービス提供に関する規則案発表、顧客の身元確認や報告を義務化」『ビジネス短信』日本貿易振興機構、2024 年 1 月 30 日。

[37] 2023 年に議会に提出された AI 関連法案は、2022 年の 88 本から 2 倍以上の 181 本となった（前掲注 30 参照）。

図表 10　AI大統領令における主要な構成要素

項目	詳細
安全性とセキュリティの新基準	NIST は、AI システムが一般公開される前のテストに厳格な基準を設定する。国土安全保障省は、これらの基準を重要インフラ分野に適用し、AI 安全保障委員会を設立する。また、国家や経済の安全保障、公衆衛生や安全性に重大なリスクをもたらす基盤モデルを開発する企業に対し、モデルのトレーニングを行う際の政府への通知、テスト結果の政府との共有を義務づける
米国民のプライバシー保護	議会に対し、全ての米国民、特に子供のプライバシー保護を強化するため、超党派のデータ・プライバシー法案を可決するよう求める。また、全米科学財団の実施する助成金事業「リサーチ・コーディネーション・ネットワーク」への資金提供を通じ、暗号ツールのような個人のプライバシーを保護する研究や技術を強化する
公平性と公民権の推進	AI アルゴリズムが司法、医療、住宅における差別を悪化させるために利用されないよう、家主、連邦政府の各種支援プログラム、連邦政府の請負業者に明確なガイダンスを提供する。また、AI に関連する公民権侵害の調査および起訴のベストプラクティスに関する研修、技術支援、政府機関との調整を通じ、アルゴリズムによる差別に対処する
消費者、患者、学生の権利保護	医療面では、AI の責任ある利用と、安価で命を救う薬剤の開発を推進する。また、米国保健福祉省は、安全プログラムの確立を通じ、AI が関与する有害、または安全でない医療行為の報告を受け、それを是正するよう行動する。教育面では、AI を活用した教育ツールを導入する教育者を支援するリソースの創出を通じ、教育を変革する AI の可能性を形作る
労働者の支援	雇用転換、労働基準、職場の公平性、安全衛生、データ収集に取り組むことで、労働者にとっての AI の害を軽減し、利益を最大化するための原則とベストプラクティスを開発する
イノベーションと競争の促進	研究者や学生が AI データにアクセスできる「全米 AI 研究リソース」の試験運用を通じ、米国全体の研究を促進する。医療や気候変動など重要分野における助成金を拡大し、米国全体の研究を促進する
外国における米国のリーダーシップの促進	国務省は商務省と協力し、国際的な枠組みを構築する取り組みを主導する。国際的なパートナーや標準化団体との重要な AI 標準の開発と実装を加速し、技術の安全性、信頼性、相互運用性を確保する
政府による AI の責任ある効果的な利用の確保	政府全体で AI 専門家の迅速な採用を加速するとともに、権利と安全を保護するための明確な基準や各省庁が AI を利用する際の明確なガイダンスを発行する

（出所）White House, “Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence,” October 30, 2023、「バイデン米政権、AIの安全性に関する新基準などの大統領令公表」『ビジネス短信』日本貿易振興機構、2023年11月1日、より野村資本市場研究所作成

3．EU

EUでは、域内発のビッグテックに相当する企業が存在せず、米国のビッグテックにAI関連市場や技術を支配されるのではないかといった危機感もあることから、他地域に先駆けて厳しい規制を指向してきた[38]。そのような中、法的拘束力を持つ世界初の包括的なAI規制法となる「AI規則」が2024年5月に成立し、同年8月に発効した[39]。

AI規則は、（1）EUの単一市場において、官民双方による安全かつ信頼できるAIシステムの開発と導入を促進、（2）EU市民の基本的権利の尊重を確保し、欧州におけるAIへの投資及びイノベーションを促進、を目的としている[40]。同規則の主な特徴としては、（1）リスクベースのアプローチ、（2）域外適用、（3）違反時の制裁金、が挙げられる。

1点目は、リスクに応じて対象のAIシステムを分類し、要件、義務及び罰則を定めるも

[38] 前掲注21、「【編集長の眼】EUで厳罰付きのAI規則法が成立、心もとない日本企業の備え」『日経クロステック』2024年5月23日。

[39] European Union, “2024/1689 Regulation (EU) 2024/1689 OF the European Parliament and of the Council of 13 June 2024; Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act),” *Official Journal of the European Union*, July 12, 2024.

[40] 駐日欧州連合代表部「EU理事会、AIに関する世界初の国際的規制を承認」2024年5月21日。

のである（図表 11 参照）。リスクが高い順に、（1）「許容できないリスク」は、基本的人権に対する侵害等の普遍的な価値に反するとされ、活用が禁止されるもの、（2）「ハイリスク」は規制対象の中心となる AI であり、既存規制下の製品（医療機器、玩具、産業機器等）と本規則で定めるものの 2 タイプが存在、（3）「特定の透明性が必要なリスク」は、人と自然相互作用する AI、感情推定や生体分類を行う AI、人物など現実世界に実在するものに酷似させたコンテンツ（ディープフェイクコンテンツ）を生成する AI 等、（4）「最小リスク」は上記のいずれにも該当しないもの、である[41]。

2 点目について、AI 規則の適用範囲は、EU 域内で上市される AI システム及び汎用目的型 AI モデルと関連するステークホルダー（プロバイダー、デプロイヤー、インポーター及びディストリビューター）とされている[42]。また、本規則は、EU 域外にも適用される。例えば、EU 域内に拠点を有さない日本企業でも、EU 域内に製品やサービスを提供する事業者は対象となる。これは、EU 一般データ保護規則（GDPR）と同様の位置付けと言える[43]。

3 点目について、AI 規則の違反時の制裁金の仕組みがあり、GDPR に比して重い水準となっている[44]（図表 12 参照）。

図表 11　AI リスクの分類

リスクレベル	概要
許容できないリスク	・ 人の生命や基本的人権に対して、直接的に脅威をもたらすと考えられる AI システム ・ 不利な扱いとされる社会的格付、基本的な行動を促す音声アシストなどが挙げられる
ハイリスク	・ 人の健康や安全、基本的人権、または社会的／経済的な利益に影響を与える可能性がある AI システム ・ 製品のセーフティコンポーネント、雇用・労働者の管理、警察や消防への緊急通報に関する AI システム等の必要となる公共・民間サービスなど
特定の透明性が必要なリスク	・ 深刻なリスクはないが、透明性に関する特定の要件を満たす必要がある AI システム ・ コンテンツや応答が、対話型 AI によって生成されたことを明らかにする要件などが挙げられる
最小リスク	・ リスクがごくわずか、またはリスクの伴わない AI システム

（注）　規則案に基づく事例。
（出所）PwC「欧州（EU）AI 規制法の解説」2024 年 9 月 9 日、より野村資本市場研究所作成

41 PwC「欧州（EU）AI 規制法の解説」2024 年 9 月 9 日。

42 汎用目的型 AI モデルは、大規模データを用いてトレーニングされた汎用性を持つ AI モデル（テキスト、音声、動画像等を作ることができる生成 AI モデルなど）。プロバイダーは、AI システムまたは汎用目的型 AI システムを開発する、開発させて上市する、または自己の名称が商標で運用する自然人、法人、公的機関等。デプロイヤーは、原則として AI システムをその権限の下で使用する自然人、法人、公的機関等。インポーターは、第三国で設立された自然人または法人の名前または商標が付いた AI システムを市場に投入する、EU に所在するまたは設立された自然人または法人。ディストリビューターは、AI システムを域内で利用できるようにする、プロバイダーまたはインポーター以外の、サプライチェーン内の自然人または法人（PwC「欧州（EU）AI 規制法の解説」2024 年 9 月 9 日）。

43 「【編集長の眼】EU で厳罰付きの AI 規則法が成立、心もとない日本企業の備え」『日経クロステック』2024 年 5 月 23 日。

44 EY 弁護士法人「欧州の AI 法規制の現状と日本企業への影響」2024 年 3 月 21 日。

図表 12　制裁金の概要

項目	内容
禁止行為違反	第 5 条に規定される AI システムに係る規制に違反した場合、最高 3,500 万ユーロまたは違反者が企業の場合は前事業年度の全世界における年間売上高の7%のいずれかの高い方の制裁金を課すものとする
AI 規則のその他の規定違反	AI システムが AI 規則の規定（第 16 条、第 22 条、第 23 条、第 24 条、第 26 条、第 31 条、第 33 条〔1〕、第 33 条〔3〕、第 33 条〔4〕、第 34 条、第 52 条）に違反している場合（第 5 条に規定されるものを除く）、最高 1,500 万ユーロまたは違反者が企業の場合は前事業年度の全世界における年間売上高の 3%のいずれか高い方の制裁金の対象とする
不正確、不完全または誤解を招く情報の提供	要請への応答として、不正確、不完全、または誤解を招くような情報を通知先機関若しくは国家の管轄当局に提供した場合、750 万ユーロ以下の罰金または違反者が企業の場合は前事業年度の全世界における年間売上高の1%のいずれか高い方の制裁金の対象とする
中小企業に対する閾値の引き下げ	中小企業（スタートアップを含む）の場合、記載されたパーセンテージまたは金額のいずれか低い方を上限とする
汎用 AI モデルのプロバイダーに対する制裁金	欧州委員会は、汎用 AI モデルのプロバイダーが故意または過失により AI 規則の適用規定に違反する等の要件に該当した際、前事業年度の世界における年間売上高の 3%若しくは 1,500 万ユーロのいずれか高い方を上限とする制裁金を課すことができる

（出所）シティユーワ法律事務所「EU AI Act の概要」2024 年 5 月 30 日、より野村資本市場研究所作成

EU における「規則」は、加盟国の国内法制化を経ない限り直接効力が生じない「指令」とは異なり、国内法制化を経なくても効力が生じるという意味で、日本法でいう「法律」に近いものと解釈されている[45]。同規則は、規制内容に応じて 2025 年 2 月から 2030 年 12 月にかけて段階的に施行される予定となっている[46]。

Ⅳ 企業による取り組み

企業による取り組みについては、（1）ISS コーポレート[47]による米国 S&P500 指数採用企業（S&P500 企業）を対象とした調査、（2）日本経済新聞社による日本の大企業の社長を対象としたアンケート調査、を取り上げる。これらの調査結果からは、企業が AI 関連リスクの高まりや政府による規制の制定、ガイドラインの公表等も背景に、AI リスクに対応する取締役会の監督体制や指針の整備、情報開示等の対応を進めている状況が浮き彫りになっている。

１．ISS コーポレートによる米国 S&P500 企業を対象とした調査

ISS コーポレートは 2024 年、2022 年 9 月～2023 年 9 月までに公表された S&P500 企業による委任状（プロキシー・ステートメント[48]、フォーム DEF 14A）に記載された AI に関連する取締役会の監督と取締役のスキルに関する記述を基に行った調査分析結果を公表した[49]。

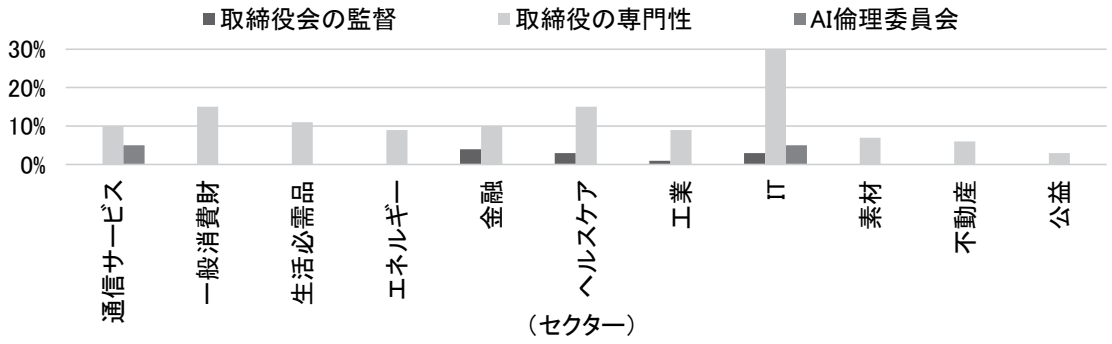
45 前掲注 3 参照。
46 前掲注 41 参照。
47 ISS コーポレートは、発行体向けサービスを担っている ISS ESG から分離・独立した米国の組織。データ、ツール、アドバイザリー・サービスを提供することにより、企業のコーポレートガバナンス、役員報酬また各企業の目標に沿ったサステナビリティ・プログラムを設計・管理し、リスクを削減し、多様な株主のニーズに対応できるよう支援している（日本取引所グループ「ESG 評価機関等の紹介　ISS ESG」）。
48 プロキシー・ステートメントは、日本の招集通知の株主総会参考書類に類似するもので、総会当日に投票される議題等が記載されており、米国証券取引委員会（SEC）に提出が義務付けられている。
49 ISS Corporate, “AI and Board of Directors Oversight: AI Governance Appears on Corporate Rader,” 2024.

調査分析結果では、取締役会による AI 関連の監督について、（1）委任状で何らかの情報開示を行っている企業は全体の約 15%、（2）情報開示を行っている企業のセクターは、情報技術（IT）が最も多く、ヘルスケア、一般消費財、通信サービスと続いている、等が示された（図表 13 参照）。

取締役会の専門性の観点からは、（1）AI に関する専門知識を有する取締役が少なくとも 1 人以上いる企業の割合は全体の約 13%で、その割合は IT セクターが最も多い、（2）AI 関連委員会や AI 倫理委員会の設置を明示している企業の割合はそれぞれ全体の約 1.6%と約 0.8%、と示された。そして、AI 関連委員会の全体像を具体的に開示している企業は全体の約 1.6%だが、その中では金融セクターがセクター全体の 4.2%と最も高い割合を占めていると記されている。

その上で、ISS コーポレートは、AI に関する明確なガイドラインや情報開示規制は調査分析結果の公表時点で存在しないものの、米国証券取引委員会（SEC）や機関投資家によるリスク管理に関するスタンスを参考として示している（図表 14 参照）。それと共に、AI が多くの企業にとって重要（マテリアル）な要素になるにつれて、投資家は、取締役会のスキルや監督責任といった要素のみならず、AI 関連全体の情報開示強化を求めていくだろうとの見方を明らかにしている。

図表 13　取締役会における AI に関する監督等に関する情報開示の状況



（注）　S&P500 指数採用企業全体に占める割合。
（出所）ISS Corporate, “AI and Board of Directors Oversight: AI Governance Appears on Corporate Rader,” 2024、より野村資本市場研究所訳

図表 14　SEC や機関投資家によるリスク管理に関するスタンス

項目	詳細
SEC	取締役の選任に関する委任状には、企業のリスク監視における取締役会の役割についての議論が含まれていなければならない。ただし、SEC は、リスク監視の運営について企業に柔軟性を与えている（例:取締役会全体、独立委員会等）
機関投資家	ブラックロック、バンガード、ステート・ストリートといった大規模な機関投資家は、重要なリスクに対する取締役会の監督について様々な方針を有している。多くの投資家は、企業は重要なリスクを軽減するためのプロセス（例：企業リスクの変化に応じて進化できる適切な監視プロセス）を確立すべきであると考えている。ただし、議決権行使の方針は、重要なリスクに関して自然な進化を許容する SEC と同様に、企業リスクを明確に定義しているわけではない

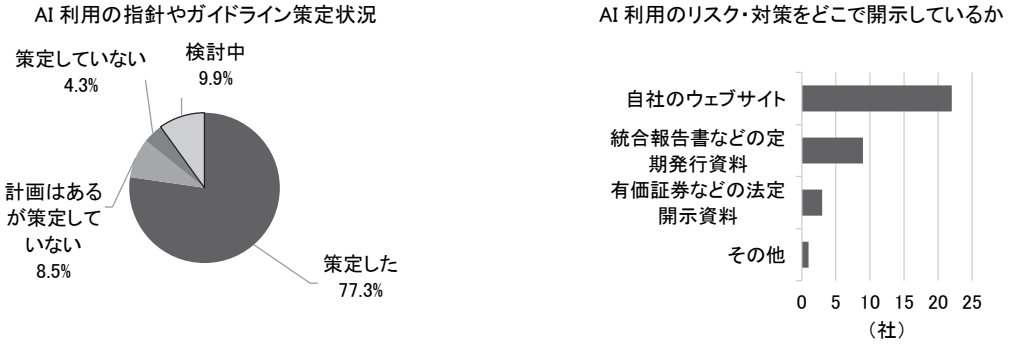
（出所）ISS Corporate, “AI and Board of Directors Oversight: AI Governance Appears on Corporate Radar,” 2024、より野村資本市場研究所訳

２．日本経済新聞社による日本企業の社長を対象としたアンケート調査

日本経済新聞社は 2024 年 4 月、同年 2 月 26 日～3 月 13 日に日本企業の社長（回答者：146 名）に対して AI に関するアンケート調査を実施した[50]。調査結果では、何らかの指針を策定している企業は有効回答企業 141 社のうち約 8 割にのぼっていることが明らかになった（図表 15 左上参照）。一方で、AI 活用の指針や想定されるリスクを情報開示している企業は 2 割弱で、大半が自社のウェブサイトに掲載しており、一部が統合報告書や有価証券報告書に記しているとのことだった（図表 15 右上参照）。また、情報開示を行っていない企業による理由は、「準備を行っているため」が最も多く、「必要性の優先順位が低い」、「AI を利用した製品、サービスを開発・提供・利用していない」といった回答もあったとのことである。

主なコメントを踏まえると、「AI 事業者ガイドライン（第 1.0 版）」や EU の AI 規則等も通じて、AI 関連リスクへの対応を喫緊の課題とする企業が多いことが明らかになっている（図表 15 下参照）。

図表 15　日本企業の社長を対象とした AI に関するアンケート調査結果



主なコメント

- 「利便性や豊かさをもたらすはずの AI が偏見や不公正な判断を助長しないよう、リスクを適正に評価し、早めに対策やルールづくりをする必要がある」
- 生成 AI は「情報漏洩、誤情報や偽情報の生成・拡散、学習データによるバイアス、権利侵害」などのリスクがある
- 「全従業員・業務に普及していく」
- 「『取り組まないことこそが最大のリスク』との論調が主流」
- 「顧客企業に AI ガバナンス構築を提案する活動もしている」
- 「EU の AI 法など各国の動向を踏まえる」
- 「AI 事業者ガイドライン（第 1.0 版）」は、「意義は大きいと感じる」一方、「政府見解は裁判規範としての法律効果はなく、AI に関するトラブルがどのように係争事案に発展し、判決が示されるのかなど海外の裁判事例に関心を持つことも必要」
- 「AI 活用が直接お客様のサービスに影響を与える際には、積極的に情報を開示していく」
- 「自社の AI に対する考え方を社内外に発信することは組織全体の共通言語をもつことでもあり、株主や投資家の期待に応えることにもつながる」

（出所）「『AI 指針』策定企業が 8 割に　富士通は倫理リスク自動抽出－社長 100 人アンケートから」『NIKKEI Digital Governance』日本経済新聞社、2024 年 4 月 18 日、より野村資本市場研究所作成

[50] 「『AI 指針』策定企業が 8 割に　富士通は倫理リスク自動抽出－社長 100 人アンケートから」『NIKKEI Digital Governance』日本経済新聞社、2024 年 4 月 18 日。

V 投資家による取り組み

投資家による取り組みについては、（1）ワールド・ベンチマーク・アライアンス（WBA）による倫理的 AI に関する投資家ステートメント、（2）米運用大手のフェデレーテッド・ハーミーズによる活動、（3）米経営コンサルティング会社の FTI コンサルティングによる AI 関連の株主提案に関する調査、を取り上げる。これらの取り組みは、投資家が企業に対して原則・ガイドラインの策定や情報開示等の対応を求めているほか、AI の普及が幅広く進んでおり、ビッグテックといった一部のセクターのみならず、より広範なセクターの企業の対応状況にも目を向け始めている現状が明らかになっている。

１．WBA による倫理的 AI に関する投資家ステートメント

ワールド・ベンチマーク・アライアンス[51]（WBA）は 2024 年 2 月、「倫理的 AI に関する投資家ステートメント（2024 年版）」[52]を公表した。WBA では 2022 年 9 月以降、「デジタル・インクルージョン[53]・ベンチマーク[54]」による評価を受けている企業（200 社）のうち、倫理的 AI 原則を開示していない 44 社と協力してきた。その結果、2023 年 9 月時点で 52 社が開示している。その一方で、投資先の企業に対して、人権の尊重と誰も取り残さないという原則に基づいて、AI に関して 4 つの要件を実装、実証し、開示することを求めているとしている（図表 16 参照）。

図表 16　WBA が企業に対して求める AI に関する 4 つの要件

- 企業のAIツールの開発、展開、調達の指針となる一連の倫理原則
- AIの開発と利用のバリューチェーン全体にわたる強力な AI ガバナンスと監督
- これらの原則が、製品やサービスのレベルを含め、企業のビジネスモデルに関連する特定のツールや行動プログラムを通じてどのように実施されているか
- AI に適用されるインパクト評価プロセス。特に高リスクのケースでは人権インパクト評価（HRIA）を重視

（出所）World Business Alliance, “2024 Investor Statement on Ethical AI,” February 26, 2024、より野村資本市場研究所訳

[51] WBA は、持続可能な開発目標（SDGs）の達成に向け、世界で最も影響力のある企業 2,000 社に説明責任を課すため、無料で一般公開されるベンチマークを開発する非営利団体（NPO）。

[52] World Business Alliance, “2024 Investor Statement on Ethical AI,” February 26, 2024.

[53] デジタル・インクルージョンとは、デジタル・ディバイド（インターネットやパソコン等の情報通信技術を利用できる者と利用できない者との間に生じる格差）を解消し、すべての人々がデジタル化の恩恵を受けられることを目指す理念。

[54] デジタル・インクルージョン・ベンチマークとは、WBA が世界で最も影響力のあるテクノロジー企業 200 社を対象に、よりインクルーシブな社会を推進する責任について評価し、ランク付けしたもの（World Business Alliance, “Digital Inclusion Benchmark,” March 13, 2023）。

２．フェデレーテッド・ハーミーズによる活動

フェデレーテッド・ハーミーズは 2018 年 4 月以降、AI に関して企業との対話（エンゲージメント）を続けており、その取り組みの多くは、効果的にリスクを軽減すべく、企業が倫理的な AI ガバナンス原則を確立できるようにすることを目的としている[55]。

同社の活動のベースとなっているのが、（1）AI とデータガバナンスに関する原則及びエンゲージメント・フレームワーク（2019 年 4 月公表）、（2）デジタル権利原則（2022 年 4 月）、となっている。

１）AI とデータガバナンスに関する原則及びエンゲージメント・フレームワーク

フェデレーテッド・ハーミーズは 2019 年 4 月、「責任ある AI とデータガバナンスに対する投資家の期待」[56]と題した論文を公表した。同論文では、投資家における環境・社会・ガバナンス（ESG）考慮事項としての AI の重要性について概説すると共に、AI とデータガバナンスに関する原則及びエンゲージメント・フレームワークを示している。

1 点目について、フェデレーテッド・ハーミーズによる「責任ある AI とデータガバナンスに関する原則」は、6 つの項目で構成されている（図表 17 参照）。同社では、投資家に対して本原則を責任ある AI の利用に適用することを薦めている。そして、上述の原則を基に策定したエンゲージメント・フレームワークは、（1）重要な法的及び財務的アウトカムを対象としたリスク要因評価、（2）顕著な社会的インパクトを対象としたプロセス主導型アプローチ、に分別されている（図表 18 参照）。

図表 17　「責任ある AI とデータガバナンスに関する原則」の概要

項目	内容
信頼	企業は、データ・プライバシーの権利についてユーザーを教育することで信頼を獲得し、完全に自由な選択肢を提供することで、ユーザーにデータの使用に対するコントロールと同意の権利を与える必要がある
透明性	企業は、バリューチェーン全体の追跡方法について透明性を確保し、データガバナンスの堅牢性と AI の公正かつ安全な使用をどのように測定しているかを開示する必要がある。企業は、スコアリングとスクリーニングの目的でデータが使用されている場合は、ユーザーに通知する必要がある
アクション	企業は、差別につながる可能性のあるデータやプロセスのバイアス等の意図しない結果を回避するために、誠実かつ徹底的に調査し、あらゆる合理的な努力を払う必要がある
誠実さ	企業は、顧客、サプライヤー、ユーザーの対応において誠実さを示す必要がある。ターゲット広告の限界を超えたショッピング、ゲーム、デバイス中毒など、中毒を助長するアプローチを含むユーザー操作を避ける必要がある。企業は、中毒に関するリスク免責事項を用意し、ターゲット広告からオプトアウトするオプションをユーザーに提供することを検討する必要がある
説明責任	企業は、AI 開発とアプリケーションのエコシステム内で、社内外に明確な説明責任体制を確立する必要がある。サプライチェーンと第三者のアクセスに関する適切なデューデリジェンスプロセスが必要である。企業は、監査可能なシステムを構築し、可能な場合は適切な保険を導入する必要がある
安全性	人間の安全は最も重要である。特に、水、電気、医療などの重要なサービスへのアクセスや、自動運転車などの交通機関の管理に関しては重要である。企業は、自社の AI アプリケーションが利益や収益よりも人間の安全を優先していることを示す必要がある

（出所）Federated Hermes, “Investors’ Expectations on Responsible Artificial Intelligence and Data Governance,” April 2019、より野村資本市場研究所訳

[55] Federated Hermes, “The Federated Hermes 2024 Outlook.”

[56] Federated Hermes, “Investors’ Expectations on Responsible Artificial Intelligence and Data Governance,” April 2019.

図表 18　「AI とデータガバナンスに関するエンゲージメント・フレームワーク」の概要

項目	内容
重要な法的及び財務的アウトカム：リスク要因アプローチ	
規制リスク	EU の「信頼性を備えた AI のための倫理ガイドライン」を始めとした、新たな法律、規制、行動規範、ガバナンス基準を認識しておく必要がある。データ・プライバシー関連問題については、EU の一般データ保護規則（GDPR）または適用されるデータ保護法(例えば、米国カリフォルニア州で制定されたプライバシー法等)を参照する必要がある
カウンターパーティ・リスク	企業は、GDPR 違反、契約違反、過失の申し立てを回避するために、データブローカーや AI 分析プロバイダーなど、供給先の取引先に対して適切なデューデリジェンスを実施する必要がある
サイバーセキュリティ・リスク	企業は、未承諾のサードパーティ・アクセスやデータ盗難を回避するために、目的に適したリスク分類とサイバーセキュリティ・アーキテクチャ（サイバーセキュリティ機能が正しく動くための全体的な構造）を確保する必要がある。英国と EU では、企業は特定のサイバーレジリエンスと違反報告法(英国のネットワーク・情報システム規則等)を認識している必要がある
悪用リスク	企業は、ビジネス上の利益を期待どおりに提供するため、AI 開発に必要な知的財産権と契約上のライセンスの確保が必要である
運用リスク	企業は、責任、クレーム、風評被害が生じる可能性のある実行リスクを徹底的に調査し、対処する必要がある
顕著な社会的インパクト：プロセス主導型アプローチ	
入力バイアス	企業は、機械学習の目的ごとに明確な根拠を持っている必要がある。特にサードパーティのデータブローカーまたはプロバイダーが関与している場合は、生成データやシミュレートされたデータを含む入力データのソース、性質、所有権を識別する強固な能力を持っている必要がある
プロセスバイアス	人間には本質的なバイアスがあるため、どのようなプロセスにも何らかのバイアスがあることを想定する必要がある。企業は、そのようなバイアスを試すことが可能な体系的アプローチを使用して、各 AI 機能のバイアスを認識する必要がある。モデリングチーム内には、可能な限り意図しないバイアスを特定するための多様性、ガバナンス及び挑戦的なプロセスが必要である
アウトカム・バイアス	データ・インプットとプロセスがアウトプットを生成するが、アウトプットの結果がアウトカムを生み出す。人間のシステムによって作成されたデータ・インプットの中には、本質的にバイアスがかかっているものがある。これには、宗教的及び政治的見解、文化的規範、ジェンダー及び民族的認識、または単純な言語の使用が含まれる。企業は、文脈によって影響を受けるデータ・インプットに基づく意図しないアウトカム・バイアスを認識し、責任ある AI の原則を適用することによって、そのような不可避なバイアスに対処する必要がある。この原則は、説明可能性と監視の 2 つの領域にまとめられている
説明可能性	説明可能性は、AI アプリケーション(インプット／プロセス／アウトプット)を理解可能なステップに分解する。特に、結果が偏っていると認識されている場合は、ユーザーとの信頼を構築することが不可欠である。透明性は、企業とステークホルダー間の情報の非対称性を軽減し、潜在的な不信感を軽減する
監視	企業は、収益や利益よりも人間の安全を優先して、AI の影響に責任を負うべきさまざまな関係者を特定する能力を備え、新たなビジネスモデルや移行中のビジネスモデルにおいて AI とデータガバナンスがもたらすリスクと機会を完全に理解する必要がある

（出所）Federated Hermes, “Investors’ Expectations on Responsible Artificial Intelligence and Data Governance,” April 2019、より野村資本市場研究所訳

２）デジタル権利原則

フェデレーテッド・ハーミーズは 2022 年 4 月、「デジタル権利原則」[57]を公表した。同原則は、AI に特化したものではない。しかし、上述の「責任ある AI とデータガバナンスに対する投資家の期待」を始めとした既存の基準等を参照した上で、情報通信技術（ICT）セクターやデータ依存度が高いセクターのためのハイレベルなエンゲージメント・フレームワークを提供していると説明されている（図表 19 参照）。特に、デジタル製品・サービスに関する負の社会的インパクトとしては、AI の誤用が挙げられ、インパクトの状況を調査するのみならず、結果について透明性を確保すると共に、規制当局に適切な権限を渡すことも必要と示している。

[57] Federated Hermes, “EOS Digital Rights Principles,” April 2022.

図表 19 「デジタル権利原則」の概要及び AI に関する説明（抜粋）

概要	
項目	投資家の期待
負の社会的インパクト	・ **AI** に対する強固なガバナンスとポリシーを確保する ・ 負の社会的影響への対処において、**子どもと若者**を優先する ・ **サプライチェーン**におけるコミュニティと労働者の権利を保護する ・ **デジタル・ディバイド**を解消するための行動をとる
表現の自由	・ **検閲**を課す法律や規制に対応するためのプロセスを維持する ・ ソーシャルメディアに透明性のある**コンテンツ・モデレーション**・ルールを実装する ・ **ネットワークの中断やシャットダウン**の命令に対応するための明確なプロセスを維持する ・ **ネットの中立性**(注 2)に関する公序良俗の立場を開示する
プライバシーの権利	・ **ユーザーに関する情報の要求**に対応するためのプロセスを維持する ・ **直接アクセス契約**に関するプロセスを維持する ・ **顔認識技術**の責任ある使用 ・ **サイバーセキュリティ**に関する堅牢なガバナンスとポリシーを確保する ・ ユーザー自身によるデータの収集、保存、および利用に関する**ユーザーの同意**を得る
AI に関する説明（抜粋）	
企業は AI に対する強固なガバナンスとポリシーを確保する必要がある。AI は、オンラインコンテンツ、ターゲット広告、検索結果、政治ニュースを選択的にまとめ、ランク付け、推奨するために、ICT セクター等で利用されている。AI は人類の発展を促進する一方、悪用される可能性もある。企業は、消費されるメディアを大幅にコントロールしながら、ユーザーの行動に影響を与えたり、社会的分断を助長したりすることを通じて、強力になる可能性がある。意図しない人種、性別、その他のバイアスがアルゴリズム内で特定されており、不公平な結果につながる可能性がある	
フェデレーテッド・ハーミーズによる「責任ある AI とデータガバナンスに関する投資家の期待」では、完全なエンゲージメント・フレームワークを提供している。企業は、アルゴリズム・システムを使用する目的の範囲を開示する必要がある。何のために最適化し、どのような変数を考慮するかを含め、どのように機能するかを説明する必要がある。加えて、ユーザーが自身の経験を形成することを決定できるようにすることが求められる。企業は、AI におけるバイアスを特定するための「EqualAI(注 3)チェックリスト」で推奨されている内容を含め、アルゴリズムにおける意図しない人種、性別、その他のバイアスを排除するためのアクションを実行する必要がある	

（注） 1. 太字は本文に基づく。
2. ネットの中立性は、インターネット上でやり取りされるデータが、コンテンツやサービスの種類によらず、常に平等に扱われること。政府やプロバイダーが、ある特定のコンテンツやサービスのデータ通信の速度や容量を、優遇または制限しないようにすること。
3. EqualAI は、AI の開発・利用に関する無意識の偏見を減らすことに取り組む米国の NPO。

（出所）Federated Hermes, “EOS Digital Rights Principles,” April 2022、より野村資本市場研究所作成

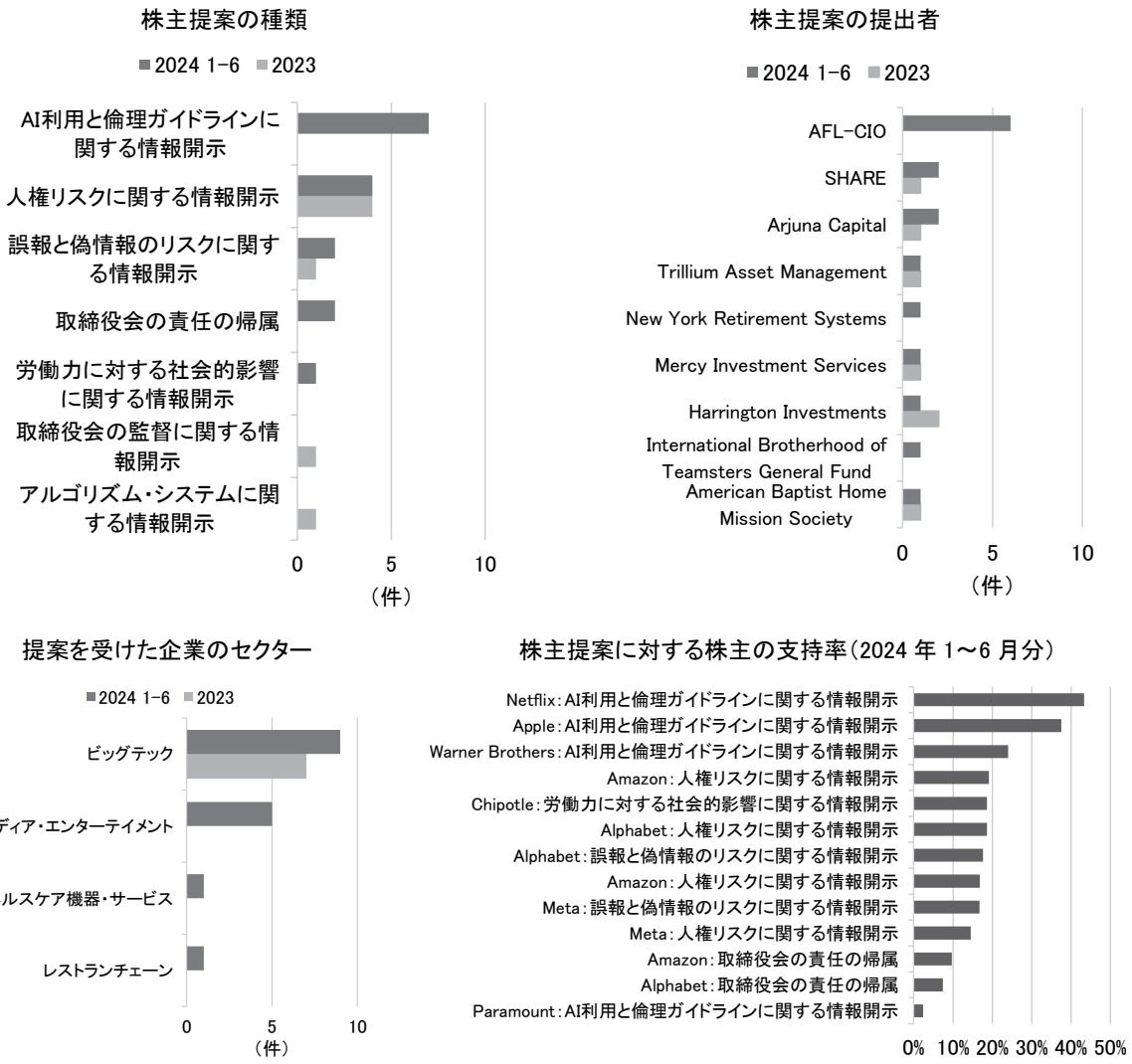
3．FTI コンサルティングによる米国の AI 関連株主提案に関する調査

FTI コンサルティングは 2024 年 9 月、米国企業の株主総会招集通知及び臨時報告書（Form 8-K）に基づき、2023 年 1 月 1 日から 2024 年 6 月 30 日の間に米国企業の株主から提出された 23 件の AI 関連の株主提案に基づく調査結果を公表した[58]。調査結果では、（1）株主提案は 2022 年の 4 件、2023 年の 7 件から 2024 年 1～6 月には 16 件と増加傾向、（2）提案の種類としては、2024 年 1～6 月は AI 利用と倫理ガイドラインに関する情報開示を求めるものが最も多かった（図表 20 左上参照）、（3）AI の社会的リスクを懸念する投資家を中心に幅広い属性の投資家が提案を行った（図表 20 右上参照）、（4）提案を受けたのはビッグテック（アルファベット、アマゾン、アップル、メタ、マイクロソフト）が大部分を占めていたが、2024 年 1～6 月にはメディア・エンターテイメント、レストラン

58 FTI Consulting, “Unveiling Key Trends in AI Shareholder Proposals,” September 9, 2024; FTI Consulting, “Increase in Shareholder Concerns Over AI Risks in Reflects Growing Investors Scrutiny, According to FTI Consulting Report,” September 9, 2024.

チェーン、ヘルスケア関連企業が初めて対象になった（図表 20 左下参照）、（5）2024 年 1～6 月の株主提案に対する支持率は全て 5 割を下回ったものの、ばらつきがある結果となった（図表 20 右下参照）、等が紹介されている。

図表 20　米国企業の株主から提出された AI 関連の株主提案に基づく調査結果



（出所）FTI Consulting, “Unveiling Key Trends in AI Shareholder Proposals,” September 9, 2024、より野村資本市場研究所作成

特に、提案を行った投資家の主な分類としては、（1）組合ファンド、（2）社会的責任投資ファンド／アドバイザー、（3）信仰に基づく投資家、（4）年金基金、が挙げられると紹介されている（図表 21 参照）。この中では、組合基金の米国労働総同盟・産業別組合会議（AFL-CIO）[59]が 2024 年 1～6 月に最も多くの提案を行ったとされている。

[59] 米国労働総同盟産業別組合会議（AFL-CIO）は、米国最大の労働組合中央組織。1955 年、AFL（米国労働総同盟）と CIO（産業別組合会議）が合同して結成。米国の組織労働者の約 8 割が加盟している（『デジタル大辞泉』小学館）。

図表 21　米国企業の株主から提出された AI 関連の株主提案の提案者別状況

提案者名	内容
組合ファンド	
米国労働総同盟・産業別組合会議（AFL-CIO）	2024 年 1～6 月の年次総会の開催期間中に最も活発な提案者となった。テクノロジー、メディア、エンターテインメント企業（アップル、コムキャスト、ディズニー、ネットフリックス、ワーナー・ブラザース）に対して、AI の利用と倫理ガイドラインに関する報告を求める提案を提出した。またアマゾンに対しても、AI に関連する人権リスクに対処するための新しい取締役会委員会の設置を求める別の提案書を提出した
物流業界労働組合連合会一般基金（International Brotherhood of Teamsters General Fund）	2024 年にテクノロジー、メディア、エンターテインメント業界以外の企業を対象とした数少ない株主の 1 つ。レストランチェーンのチポトレに対し、AI の利用と自動化に伴う労働力への影響に対処する計画について情報開示するよう求めた
社会的責任投資ファンド／アドバイザー	
アルジュナ・キャピタル（Arjuna Capital）	2023 年と 2024 年に、ビッグテック（アルファベット、メタ、マイクロソフト）に対して、AI が生成する誤情報や偽情報のリスクに関する情報開示を求める提案書を提出した
シェアホルダー・アソシエーション・フォー・リサーチ・アンド・エデュケーション（SHARE）	2023 年と 2024 年に、アップルに対して、ターゲティング広告の実施に伴う人権リスクに関する報告を求める提案書を提出した。2024 年には、ユナイテッドヘルス・グループに対しても、AI の使用と監督に関する追加情報の開示を求める提案書を提出した
トリリアム・アセット・マネジメント（Trillium Asset Management）	2022 年と 2023 年にアルファベットに対して、アルゴリズム・システムの透明性に関する情報開示を求める提案書を提出した。2024 年になって、エンゲージメントをエスカレーションし、監査コンプライアンス委員会の憲章を更新し、AI に関する責任を反映するよう求める提案書を提出した
ハリントン・インベストメンツ（Harrington Investments）	アマゾンに対して過去 6 年間、同じ提案書を提出している。具体的には、同社の顔認識技術の使用に関する人権リスクについて、第三者による報告を求めている。2023 年にはアルファベットに対しても提案書を提出し、AI 利用に関連するものを含め、同社の業務からステークホルダーにもたらされ得る重大なリスクを効果的に監督する上での監査コンプライアンス委員会の役割について、独立した評価を求めている
信仰に基づく投資家	
米国バプテスト・ホーム伝道協会（American Baptist Home Mission Society）	アマゾンに対して過去 5 年間、同じ提案書を提出している。具体的には、顧客による同社製品の使用が人権侵害につながる大規模な監視を可能にするかどうかを評価する独立した調査を求めている
マーシー・インベストメント・サービス（Mercy Investment Services）	メタに対して過去 3 年間、同じ提案書を提出している。具体的には、同社のターゲティング広告が人権に及ぼす影響に関する第三者による報告を求めている
年金基金	
ニューヨーク市退職年金基金（New York City Retirement Systems）	2024 年、パラマウントに対して、同社の AI 利用、それを監督する取締役会の役割、倫理ガイドラインに関する情報開示を求める提案を行った

（出所）FTI Consulting, “Unveiling Key Trends in AI Shareholder Proposals,” September 9, 2024、より野村資本市場研究所訳

FTI コンサルティングでは、調査結果を踏まえて、（1）2024 年 1～6 月にはビッグテック以外の企業にも株主提案が行われたことを踏まえると、2025 年にはさらに幅広いセクターで提案が行われると想定、（2）アップルやネットフリックスの投票結果等をきっかけに、組合ファンド等が情報開示に関して他の企業に株主提案を提出する可能性、といった見通しを明らかにしている。

VI 今後の論点

AI が急速に発展・普及する中で、リスクを管理・抑制しつつ最大限の便益を得ることを目的とした AI ガバナンスの重要性が世界的に認識されてきた。各国・地域でスタンスは異なるものの、政府は法規制・ガイドライン等の対応を進め、企業自身も取締役会の監督体制や指針の整備、情報開示等の対応を進めている。さらに、近年は投資家も投資先企業の AI 関連リスクに着目し、原則・ガイドラインの策定や情報開示等の対応を求めてい

るほか、ビッグテックのみならず幅広いセクターの企業の対応状況に着目しつつある。

金融資本市場の観点から、AI ガバナンスを企業等にさらに浸透させていくことは、企業価値維持・向上、ひいてはリスク調整後の投資パフォーマンスの維持・向上にもつながり得るため、重要と言える。その意味での今後の論点としては、（1）AI ガバナンスの重要性に関する理解の促進、（2）情報・データの適切な開示、（3）第三者評価等の仕組みの整備、が挙げられる。

1 点目について、本稿で紹介したフェデレーテッド・ハーミーズのように企業に対して AI に関するエンゲージメントを行う投資家が近年見られるようになっている。投資家が企業に対してエンゲージメントを実施するに当たっては、AI が IT 面のみの課題ではなく、企業価値に影響を及ぼし得る重要な経営課題であることを繰り返し伝え、企業の経営層における意識向上を促すことが大切と言える。また、エンゲージメントを行う際にはセクターによってリスクの度合いが異なるため、企業の属性や置かれた状況を踏まえた上で丁寧に対話を行うことがカギになる。例えば、ESG 評価機関の ISS ESG では、製品やサービスにおける AI 利用がより高いリスクをもたらす可能性のある 21 のセクター／機能を特定している（図表 22 参照）。

図表 22　AI 導入リスクが高いセクター／機能

セクター／機能	潜在的なネガティブ・インパクト		
	安全性	プライバシー	差別
航空宇宙／防衛	○	―	―
自動車	○	○	―
商業銀行／資本市場関連	―	○	○
デジタル・ファイナンス／支払いプロセス	―	○	○
教育関連サービス	―	○	○
電力会社	○	○	―
電子デバイス／機器	○	○	―
ガス及び電力ネットワーク事業者	○	○	―
ヘルスケア関連機器及び消耗品	○	○	―
ヘルスケア関連施設及びサービス	○	○	―
ヘルスケア関連技術及びサービス	○	○	―
人事及び雇用サービス	―	○	○
保険	―	○	○
インタラクティブ・メディアサービス／オンライン消費者サービス	○	○	―
住宅ローン及び公的セクター金融	―	○	○
複数インフラサービス	○	○	―
石油・ガス貯蔵／パイプライン	○	―	―
公的／地域金融機関	―	○	○
ソフトウェア／多様な IT サービス	―	○	―
通信	―	○	―
水道・廃棄物処理施設	○	―	―

（出所）ISS ESG, “The Intersection of Artificial Intelligence & ESG,” May 2024、より野村資本市場研究所訳

2 点目について、企業は、投資家を始めとしたステークホルダーが必要とする情報・データを適切に情報開示することが大切であることは言うまでもない。現時点では、投資の観点から企業の AI ガバナンスを評価する際に有効な定性面・定量面の情報について、世界的なコンセンサスは存在しないとみられる。しかし、前述のとおり、一部の投資家が既に AI ガバナンスをめぐってエンゲージメントや株主提案といった行動を起こし始めていることを踏まえると、徐々にイメージが明らかになってくるとみられる。その際には、投資家が情報の比較可能性を求めると考えられ、個々の企業による取り組みに加え、政府等による情報開示フレームワークやガイドラインの策定に関する検討が必要となると想定される。

3 点目について、AI は比較的新しい分野かつ技術、法務、倫理など、多様な側面に影響を及ぼし得るものであり、投資家においても企業の AI リスクやガバナンスを評価できる人材は現時点で多くないとみられる。その一方、例えば AI と同様に比較的新しい分野であるサイバーについては、第三者評価としてサイバーリスク格付け[60]が存在し、投資家が投資の意思決定等の用途で活用し始めている[61]。このような動きを踏まえると、AI に関しても格付けやスコアリングといった第三者評価がインフラとして整備されれば、投資家がより AI リスクやガバナンスを評価、分析し、投資行動に活かしやすくなるとともに、企業への AI ガバナンスの浸透を後押しする可能性もあると考えられる。

[60] サイバーリスク格付けは、企業等のセキュリティ対策を含むサイバーセキュリティ・パフォーマンスをデータに基づいて定量的に評価し、その結果を簡単な記号で表したものである。サイバーリスク格付では、企業等のインターネット接続に関する外部から観察可能なデータを使用して、いくつかのセキュリティリスク要素に基づいて企業のサイバーセキュリティ態勢について評価している。米国のセキュリティスコアカードとビットサイトが市場のリーダー的な存在とされている（牛窪賢一「米国を中心とするサイバー保険市場の動向」『損保総研レポート』第 138 号、損害保険事業総合研究所、2022 年 2 月）。

[61] 米国のサイバーリスク格付会社であるビットサイトによると、投資家は、（1）デューデリジェンスと引受の強化（サイバーリスクを投資の意思決定に反映）、（2）アルファ・シグナル（アウトパフォームの機会を特定）、（3）ポートフォリオリスク管理（リスクの高い企業を特定し、リスク軽減とエンゲージメント活動を優先）、（4）投資管理（効果的なサイバーリスクの監視を通じた長期的な投資価値の保全・向上）、等の用途で活用している（ニコル・パスター・マツセク「サイバーセキュリティと財務パフォーマンス／リスクとの重要な関連性」『野村サステナビリティクォータリー』第 5 巻第 3 号〔2024 年夏号〕）。