Marubeni
丸紅経済研究所
2024 年 5 月 3 日

# AI 規制の夜明け：国際社会が進める法的枠組みの構築とその未来

チーフ・アナリスト　重吉　玄徳
h-shige@marubeni.com

- AI は、生産性の向上やイノベーションの促進において重要な役割を果たす一方で、偽情報の拡散、意見操作、知的財産権の侵害、プライバシー侵害などの問題を引き起こすリスクも内包する。これらの課題に対応するため、国際的に AI に関するルールメイキングが進められている。
- 2023 年、広島で開催された G7 広島 AI サミットを契機に、AI の国際ガバナンスの形成を促進する目的で「広島 AI プロセス包括的政策枠組み」という初の国際的枠組みが策定され、米国では、バイデン政権による AI の安全、安心で信頼できる開発と利用に関する大統領令が発出された。欧州では、2024 年 3 月に世界初となる AI 法が欧州議会で採択され、日本では、2024 年中の AI 推進基本法の法案提出が目指されている。
- AI に関する法規制は現在、その整備に向けた初期段階にあるが、欧州での AI 法の採択や米国の大統領令の発出といった具体的な進展が見られ始めている。今後、AI の社会的影響が拡大するにつれ、プライバシー保護など、実際のビジネス環境に影響を及ぼす規則が導入されることが予想される。したがって、これらの動向を適切に把握し対応することが求められる。

人工知能（AI）の未来は、今や科学技術の最前線を超え、私たちの社会、経済、さらには日常生活における根幹を揺るがす存在となりつつある。その進化の速度と影響の大きさは、国境を越え、国際的な対話と協力を必要としている。この革新的な技術が持つ可能性を最大限に引き出しつつ、リスクを適切に管理することが極めて重要である。2023 年、世界はこの問いに答えるべく、さまざまな動きを見せた。本稿では、その動向について考察する。

## １．国際的な動向

（１）AI の国際ルールを形作る：広島 AI プロセス

近年、AI 技術の急速な発展と普及が国際社会共通の重要課題となっている。この問題を受けて、2023 年 5 月に広島で開催された G7 サミットの一環として、G7 広島 AI サミットが開催された。このサミットで、AI に関する国際的なルールを検討する「広島 AI プロセス」という新しい取り組みが始動した。G7 メンバー及び関係する国際機関が参加したこのプロセスは、12 月の閣僚級会合で、安全、安心で信頼できる高度な AI システムの普及を目的とした初の国際的枠組み「広島 AI プロセス包括的政策枠組み」を策定し、G7 首脳によって承認された。

この枠組みの目的は、民主主義国家が共有する価値観に基づいて、先進的な AI システムの開発を推進し、責任あるイノベーションと新興技術のガバナンスをリードすることにある。枠組みには、(1)OECD の報告書、(2)国際指針、(3)国際行動規範、(4)AI に関するプロジェクトベースの協力、の 4 つの要素が含まれており、これらは相互に補完しあい、AI の未来を形作るための重要な基盤となる。特に、（２）の「全ての AI 関係者向けの広島プロセス国際指針」は、 AI 関係者全体に求められる具体的な指針として位置づけられ、各国の規制策定の基準となることが期待される。各国がこの指針に基づいて規制を整備することで、AI 技術の適用における国境を越えた一貫性と互換性が保たれ、国際的な協力と調和が促進されるだろう。

| 全ての AI 関係者向けの広島プロセス国際指針 | |
|---|---|
| 1 | AI ライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度な AI システムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じる |
| 2 | 市場投入を含む導入後、脆弱性、及び必要に応じて悪用されたインシデントやパターンを特定し、緩和する |
| 3 | 高度な AI システムの能力、限界、適切・不適切な使用領域を公表し、十分な透明性の確保を支援することで、アカウンタビリティの向上に貢献する |
| 4 | 産業界、政府、市民社会、学界を含む、高度な AI システムを開発する組織間での責任ある情報共有とインシデントの報告に向けて取り組む |
| 5 | 特に高度な AI システム開発者に向けた、個人情報保護方針及び緩和策を含む、リスクベースのアプローチに基づく AI ガバナンス及びリスク管理方針を策定し、実施し、開示する |
| 6 | AI のライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ、内部脅威に対する安全対策を含む、強固なセキュリティ管理に投資し、実施する |
| 7 | 技術的に可能な場合は、電子透かしやその他の技術等、ユーザーが AI が生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入する |
| 8 | 社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策への投資を優先する |
| 9 | 世界の最大の課題、特に気候危機、世界保健、教育等（ただしこれらに限定されない）に対処するため、高度な AI システムの開発を優先する |
| 10 | 国際的な技術規格の開発を推進し、適切な場合にはその採用を推進する |
| 11 | 適切なデータインプット対策を実施し、個人データ及び知的財産を保護する |
| 12 | 高度な AI システムの信頼でき責任ある利用を促進し、貢献する。 |

（出所）総務省「広島 AI プロセス」より丸紅経済研究所作成

## 2．米国の動向

（1）AI の未来をリードする：米国 AI イニシアティブ

2019 年 2 月 11 日、トランプ前大統領は「AI における米国のリーダーシップの維持」に関する大統領令に署名した。この大統領令は、米国が AI 技術の開発と利用においてグローバルリーダーの地位を維持し、さらに強化するための戦略を確立することを目的としている。この大統領令に基づき開始された「米国 AI イニシアティブ」は、AI の技術革新を推進し、その安全なテストと展開に対する障壁を減らすこと、AI 時代に適した労働者の育成、そして米国が技術的優位を維持できるような国際環境の促進を目指している。このイニシアティブは、AI 技術に対する国民の信頼と信用の醸成、市民の自由、プライバシー、米国の価値観を保護しなければならないとしている。

（2）技術進化の中で保護される権利：AI 権利章典の青写真が目指すもの

2022 年 10 月に、科学技術政策局は「AI 権利章典の青写真」を公表した。この青写真は、1791 年に批准された米国の権利章典の精神を継承しつつ、21 世紀の技術進化に伴う新たな課題に対応することを目指している。元々、権利章典が 18 世紀の政治的文脈に根ざし、主に政府による権利侵害からの保護に焦点を当てていたのに対し、AI 権利章典の青写真は、AI など自動化システムによる潜在的な権利侵害への対処を企図している。この取り組みは、技術の急速な進歩がもたらす新たな課題に対応し、21 世紀の市民権を守るための重要な一歩となる。

公表された AI 権利章典の青写真は、AI 技術に接する可能性のある全米国市民の権利保護を目指している。対象は、消費者、労働者、患者、学生など、多岐にわたる。また、AI 技術の開発者、導入者、政策立案者に対しても、倫理的かつ責任ある技術の利用を促すことを目的としている。この青写真は、５つの原則から構成され、AI 技術の設計、開発、導入、利用における倫理的な枠組みとガイドラインを提供する一方で、既存の法令や規則を修正したりするものではなく、法的拘束力はない。

| **AI 権利章典の青写真　５つの原則** | |
|---|---|
| 安全で効果的なシステム | AI システムは、安全かつ効果的であることが求められる。開発過程は、多様な専門家と協力を得て進められ、リスクの特定と軽減、安全かつ効果的であることを確認するための継続的な監視を受けるべきである。 |
| アルゴリズムによる差別からの保護 | アルゴリズムによる不公平や差別を防ぐための措置が必要である。アルゴリズムによる差別から個人およびコミュニティを保護し、システムの利用と設計は公平な方法で行われるべきである。 |
| データのプライバシー | 個人のプライバシーが尊重され、厳密に必要な範囲でのみデータを収集するべきである。個人データの収集、使用、共有に関しては個人のコントロールが保証されるべきである。 |
| 通知と説明 | 自動化システムによって影響を受ける個人は、その事実とシステムの動作原理について適切に通知され、説明を受ける権利を有する。 |
| 人による代替手段、配慮、予備的措置 | 自動化システムにはオプトアウトの選択肢があり、発生した問題を迅速に検討し、解決できる担当者へのアクセスが可能であるべきである。 |

（出所）ホワイトハウス HP より丸紅経済研究所作成

（3）信頼できる AI への道：業界大手のボランタリーコミットメンツ

2023 年 7 月 21 日、ホワイトハウスは、AI の安全性や個人情報保護への懸念に応える形で、Amazon、Anthropic、Google、Inflection、Meta、Microsoft、OpenAI を含む主要な AI 企業 7 社が、AI 技術の安全で信頼できる、透明性のある開発に向けたボランタリーコミットメンツに合意したと発表した。9 月 12 日には、Adobe、Cohere、IBM、Nvidia、Palantir、Salesforce、Scale AI、Stability の追加 8 社がこのコミットメンツに合意した。

| ボランタリーコミットメンツ | |
|---|---|
| 製品の公開前に安全性を確認する | AI システムの公開前に内部および外部のセキュリティテストを実施する |
| | AI リスクの管理に関する情報を産業界及び政府、市民社会、学術界と共有する |
| 安全第一のシステム構築を行う | AI システムの中で最も重要な部分である非公開のモデルのウエイトを保護するために、サイバーセキュリティと内部脅威対策に投資する |
| | AI システムの脆弱性を第三者が発見し、報告することを容易に行えるようにする |
| 国民の信頼を得る | AI によって生成したコンテンツであることをユーザーが認識できるように、堅牢な技術メカニズムの開発に取り組む |
| | AI システムの能力、限界、適切なおよび不適切な使用領域に関する報告を公開する |
| | 有害な偏見や差別の回避、プライバシーの保護など、AI システムがもたらしうる社会的リスクに関する研究を優先する |
| | 社会が直面する最大の課題に対処するために、先進的な AI システムを開発、導入する |

（出所）ホワイトハウス HP より丸紅経済研究所作成

（4）安全で信頼できる AI：バイデン政権が示す道筋

2023 年 10 月 30 日、バイデン大統領は、主要企業 15 社によるボランタリーコミットメンツを土台として、「AI の安全、安心で信頼できる開発と利用」に関する大統領令を発令した。バイデン政権は、AI の開発と利用を安全かつ責任を持って管理することを最重要課題としており、そのために連邦政府全体で協調したアプローチを進めている。この大統領令は、これまでのボランタリーコミットメンツなどとは異なり、AI 技術の安全で信頼できる開発と利用に関する一定の措置を連邦政府機関に対して指示するものとなっており、今後、義務や罰則などが規定される可能性がある。実際、2024 年 1 月 29 日に公表された商務省の規則案では、米国のインフラストラクチャー・アズ・ア・サービス（IaaS）提供者及びそのサービスの外国再販業者に対し、外国顧客の身元情報を保持するよう義務付けられている。この規則案は、顧客特定プログラム（CIP）の維持、サイバー活動への悪用防止、および特定の AI モデル利用に関する取引の報告義務などを定め、これらの義務違反に対しては、罰則の対象となり得ることが示されている。

| AI の安全、安心で信頼できる開発と利用に関する大統領令の主な内容 | |
|---|---|
| AI 技術の安全性とセキュリティの確保 | AI の安全性とセキュリティに関するガイドライン、標準、およびベストプラクティスの策定を命じている。AI システムの評価、セキュリティリスク対処、デュアルユース基盤モデルの開発者による |

| | |
|---|---|
| | 連邦政府への情報共有の義務化、AI 生成コンテンツを識別するためのラベリングおよびコンテンツ認証に関するガイダンスの策定なども含まれる。 |
| イノベーションと競争の促進 | AI 人材を米国に誘致し、イノベーションと市場競争の促進を奨励する措置を講じる。これには、知的財産の保護と中小企業及び起業家へのイノベーション機会の提供などが含まれる。 |
| 労働者の支援 | AI が労働市場に及ぼす影響に関する報告書の作成、AI が従業員の幸福を増進させるための原則とベストプラクティスの策定、AI 対応の多様な労働力の育成などが含まれる。 |
| 公平性と市民権の推進 | 公平性と市民権を強化するために AI を使用する意向を示している。これには、刑事司法制度、公的給付、雇用、住宅市場、消費者金融市場における差別への対処などが含まれる。 |
| 消費者、患者、乗客、学生の保護 | AI の安全かつ責任ある使用を促進し、消費者、患者、乗客、学生を保護する措置が指示されている。詐欺、差別、プライバシー侵害対策、教育分野での AI 使用方針やガイダンスの策定などが含まれる。 |
| プライバシーの保護 | AI によって増大するプライバシーリスクの軽減策が提案されている。これには、AI による個人情報収集や個人に関する推論を行うことへの対処、プライバシー拡張技術の開発と実装を支援することが強調されている。 |
| 連邦政府による AI 利用の促進 | AI の効果的な使用を促進し、AI の開発および利用を調整するための省庁間協議会の設立、AI のイノベーションを推進しリスクを管理するための指針の発行、各省庁におけるチーフ AI オフィサーの指名などが含まれる。また、連邦政府内の AI 利用の透明性を高めるため、行政管理予算局長によって連邦政府における AI の使用事例の収集、報告、公開のための年次報告指示が行われる。 |
| 海外における米国のリーダーシップ強化 | AI の潜在能力を引き出し、その課題を克服するための世界的な取り組みにおける米国のリーダーシップ強化を目指している。これには、国際的な同盟国やパートナーとの協力拡大、AI のリスクを管理し利益を活用するための強力な国際的枠組みの確立、AI に関するグローバルな技術標準の推進などが含まれる。 |

（出所）ホワイトハウス HP より丸紅経済研究所作成

また、国防生産法に基づき最も強力な AI システムを開発する企業には安全性試験の結果を報告することが義務付けられ、海外で AI トレーニングを行う米国のクラウド企業に対しても報告が求められる規制が提案されている。さらに、全米国立科学財団が運営する AI 研究リソースの試験運用の開始と、データサイエンティストを含む AI 専門家の大規模採用を通じて、連邦政府全体の AI 専門家の採用を加速する「AI タレントサージ」の取り組みが進められている。連邦政府による AI 関連の動きはこれだけにとどまらず、2024 年 4 月 27 日には国防、移民、著作権など多岐にわたる項目を含む大統領令の期限が設定されており、政府の AI に関する取り組みは緒についたばかりである。

（5）新法案：Artificial Intelligence Advancement Act of 2023

政府と並行して、連邦議会もAIの規制を推進している。2023年10月17日、「Artificial Intelligence Advancement Act of 2023」というタイトルの法案（S.3050）が提出された。この法案の目的は、人工知能の発展を促進し、適切な規制枠組みを整備することである。具体的には、金融サービス業界で AI がどのような業務に頻繁に活用されているか、AI の利用に関して設けられている現行のガバナンス基準などに関する報告書の提出、国防総省が任務や作戦に組み込む AI モデルに関するバグ報奨金プログラムの開発、などが含まれている。この法案は、立法プロセスの初期段階にあり、委員会での審議を経て上院本会議へと進む可能性がある。現在、米国では AI に関する規制が積極的に議論されており、AI 技術の急速な進展に対応するための規制が整備されつつあると言える。

## 3．欧州の動向

（1）EU が目指すデジタル社会：Shaping Europe's Digital Future

欧州委員会が 2020 年 2 月に公表した「Shaping Europe's Digital Future」は、ヨーロッパのデジタル化推進とその未来形成に向けた包括的なアプローチを示している。この戦略は、EU の基本的価値観を重視し、戦略的自律を目指しており、主要な 3 つの柱は、人々のためのテクノロジー、公正で競争力あるデジタル経済、そして開かれた民主的で持続可能な社会である。

「Shaping Europe's Digital Future」の一部として公表された AI 白書では、AI 技術の発展と利用を促進するために、「卓越のエコシステム」と「信頼のエコシステム」の二つの政策枠組みを提案している。前者は研究とイノベーションを通じて AI ソリューションの採用を加速させ、後者は高リスクな AI システムに関する基本的権利と消費者権利の保護に焦点を当てている。そして、AI に対する新しい規制枠組みが、企業にとって過度に細かい規制を生じさせることなくバランスを取れたものとなるようリスクベースのアプローチを採るべきだと指摘している。

（2）AI の未来を形作る：AI 法

この流れを受け、AI の使用に関する欧州全体の規制枠組みを定める AI に関する規則案が 2021 年 4 月 21 日に公表された。2023 年 12 月 9 日には EU の 3 機関で政治的な合意が成立し、2024 年 3 月 13 日に欧州議会が本会議でこの法案を正式に採択した。

この AI 法は、世界で初めて AI 技術に対する全面的な規制枠組みを設けるものであり、歴史的な一歩を踏み出している。この法律は、単にリスクを管理するための規制を設けることにとどまらず、AI 技術が社会にもたらす利益を最大化しつつ、潜在的な害を最小限に抑えるバランスを見つけることを目指している。この法律により、AI システムの開発者や提供者は、そのシステムが EU の基本的権利や価値観を尊重し、安全であることを保証する責任が課される。これにより、消費者や市民の信頼を得ることができ、AI 技術の健全な発展を促進することが期待される。この法律は、AI 技術の将来の発展方向性を示す指標となり、他の国や地域の政策形成に影響を与える可能性がある。

主なアプローチは、社会に害を及ぼす能力に基づいて AI を規制するリスクベースに基づくもので、AI がもたらすリスクの程度に応じて、異なるレベルの規制要件が適用される。

| リスクベースアプローチ | | |
|---|---|---|
| **リスクの程度** | **具体例** | **AI の利用** |
| 容認できないリスク | サブリミナル技術を使用する AI システム、年齢、身体的もしくは精神的障害に起因する脆弱性を悪用する AI システム、自然人の信頼性を評価又は分類し不利な扱いをもたらす社会的スコア　等 | 禁止 |
| ハイリスク | 特定の EU 法の対象となる製品の安全部品として、または AI 自体が対象製品であり、その製品が第三者の適合性評価を受ける必要がある AI システム（例：玩具、無線機器、体外診断医療機器、民間航空セキュリティ、農業車両）<br>生体認証、重要インフラ、教育および職業訓練、雇用・労務管理、医療サービスを含む必要不可欠な公的扶助の給付やサービス、法執行、移民・亡命・国境管理、司法及び民主的手続きの運営　等 | 厳格な規制 |
| 限定的なリスク | 自然人とやり取りする AI、ディープフェイク生成 AI、感情認識システム　等 | 透明性の義務 |
| 最小限のリスク | 上記以外の AI システム | 新たな義務なし |

（出所）欧州委員会 HP より丸紅経済研究所作成

容認できないリスクを持つ AI システムは、社会にとって受け入れがたい AI であり、使用が禁止される。例えば、個人の意識下での行動を操るサブリミナル技術や、年齢、身体的もしくは精神的障害に起因する脆弱性を悪用する AI システムがこれに該当する。

ハイリスクな AI システムに対しては、厳格な規制が適用される。具体的には、特定の EU 法規によって安全性が要求される製品に組み込まれる AI システムや、公共サービスの提供に影響を与えるシステムなどが含まれる。これらの AI システムは、リスク管理システムの確立、品質基準に適合するデータセットを用いた開発、技術文書の作成及び更新、結果のトレーサビリティを確保するためのログ管理、透明性の確保、人間による監視、そして高水準の正確性、堅牢性、サイバーセキュリティの確保など、一連の厳格な用件が課される。

限定的なリスクの AI システムには、透明性の義務が課される。例えば、ユーザーが AI と対話するチャットボットや、人間の感情を認識するシステムなどが該当する。これらの AI システムは、人が AI とやりとりしている際にそのことを認識できるよう、AI システムの設計と開発を行う必要がある。特に、合成音声や画像、映像、テキストを生成するシステムでは、その出力が AI によって生成または操作されたものであることが認識可能であることを保証する必要がある。感情認識や生体認証システムでは、対象者への通知と適切な個人データの処理を行う必要がある。ディープフェイクや人為的に生成されたコンテンツを扱うシステムでは、その事実を開示する義務がある。

これらの AI システム自体への要求事項に加えて、運用に関する要求事項として、ハイリスクな AI の提供者の義務と利用者の義務が規定されている。ハイリスクな AI の提供者の義務には、品質管理システムの導入、技術文書の作成・保管、AI システムによって自動的に生成されるログの保管、是正

措置及び情報提供、管理当局への協力、EU 適合宣言書の作成、適合の CE マークの貼付、重大インシデントの報告などがある。AI 利用者の義務は、使用説明書に従って AI システムを使用する、AI システムを使用する際、人間による監視を確保する、入力データが適切かつ十分に代表的であることを確保する、重大なインシデントが確認された場合は関係者に通知し使用を停止する、ログを保管する、職場での使用に先立ち影響を受ける従業員に通知する、基本的権利に対する影響の評価とその結果の通知などがある。

生成 AI などの汎用目的 AI モデルに関しては、広範な用途に使用されるため、提供者の義務として透明性に関する要件が定められている。これには、AI によって生成される音声、画像、ビデオ、テキストなどのコンテンツが、人工的に生成されたか、あるいは加工されたものであることをユーザーが識別できるようにする義務が含まれる。これは、偽情報の拡散を防ぐための重要なステップになる。汎用目的 AI モデルの提供者には、学習及びテスト過程並びに評価結果を含む技術文書を作成し、最新の状態に維持すること、EU 著作権法を尊重する方針を導入すること、AI オフィスが提供するテンプレートに従って、学習に使用されるコンテンツに関する十分に詳細な要約を作成し、公開すること、欧州委員会および各国の主管庁が本規則に基づく権限および権限を行使する際に、必要に応じて協力しなければならないといった義務もある。また、多くの人々に大規模な影響を与える可能性があるシステミック・リスクを持つ汎用目的 AI モデルには、追加の義務がある。リスクを特定して軽減するための敵対的テストを実施し記録する、汎用目的 AI モデルの使用から生じる可能性のあるシステミック・リスクを評価し軽減する、AI オフィスおよび必要に応じて国の適切な当局に重要な情報を遅滞なく追跡、記録、報告するなどである。

AI 法では、違反に対する罰則も定められている。AI の使用禁止違反の場合、罰金は 3,500 万ユーロ以下、または違反者が企業である場合は、前会計年度の全世界における年間売上高の 7%以下のいずれか高い方が科される。AI システムが特定の規定を遵守していない場合、1,500 万ユーロ以下、または違反者が企業である場合、その前会計年度の全世界における年間売上高の 3%以下のいずれか高い方の罰金が科される。届出機関及び国の管轄当局に不正確、不完全または誤解を招く情報を提供した場合の罰金は、750 万ユーロ以下、または違反者が企業である場合にはその前会計年度の全世界における年間売上高の 1%以下のいずれか高い方の罰金が科される。中小企業には、上記のいずれか低い方の罰金が適用される。

今後のスケジュールについては、以下のタイムラインに沿って進められる予定である。

| **AI 法のタイムライン** | |
|---|---|
| 発効日 | EU 官報に記載された 20 日後に発効 |
| 施行日 | 特定の規程を除き、発効から 24 カ月後に施行 |
| 施行 6 ヵ月後 | 容認できないリスクを持つ AI を禁止 |
| 施行 12 ヵ月後 | 汎用目的 AI モデルの提供者に対する義務が発効<br>各 EU 加盟国内で AI 法案の実施、監視、および遵守を担当する機関や当局を指定<br>禁止された AI のリストの年次レビューおよび修正 |
| 施行 18 ヵ月後 | 上市後の AI に対する継続的な監視プロセスを施行 |

| | |
|---|---|
| 施行 24 ヵ月後 | 附属書 III に具体的に記載されたハイリスク AI システム（生体認証、重要インフラ、教育、雇用、公共サービスへのアクセス、法執行、移民および司法の運営に関する AI システムが含まれる）に対する義務が発効<br>加盟国が、罰金を含む罰則ルールを施行<br>加盟国当局が、少なくとも 1 つの運用可能な AI 規制サンドボックスを設立<br>ハイリスク AI システムのリストのレビューおよび修正 |
| 施行 36 ヵ月後 | 附属書 III に記載されていないものの、製品の安全部品として、または AI 自体が製品であり、その製品が特定の EU 法に基づいて第三者の適合性評価を受ける必要があるハイリスク AI システム（例：おもちゃ、無線機器、体外診断医療機器、民間航空セキュリティ、農業車両）に対する義務が発効 |

（出所）欧州委員会 HP より丸紅経済研究所作成

## ４．日本の動向

（１）信頼できる AI への道：AI セーフティ・インスティテュート

2024 年 2 月 14 日、政府は AI の安全性に対する国際的な関心の増加を背景に、AI セーフティ・インスティテュート（AISI）の設立を発表した。この新設された機関は、安全かつ信頼性の高い AI を実現するため、内閣府や関係省庁、関係機関との連携の下、独立行政法人情報処理推進機構に設置され、AI の安全性を確保するための基準策定と、評価手順の整備を行う。具体的には、安全性に関する調査、基準の策定、評価手法の開発、さらには国際的な協力関係の構築を主な業務として担当する。

| **AISI の業務内容** | | **2024 年 2 月 29 日時点での予定** |
|---|---|---|
| 安全性評価に係る調査、基準等の検討 | 安全性に係る標準、チェックツール、偽情報対策技術、 AI とサイバーセキュリティに関する調査 | ４月目途に各種調査事業の方針を定める |
| | 安全性に係る基準、ガイダンス等の検討 | ３月末目途にリスクマネジメントフレームワーク(RMF)の 和訳を公表する<br>５月末目途に AI 事業者ガイドラインと RMF との Crosswalk を公表する |
| | 上記に関する AI のテスト環境の検討 | ８月目途にレッドチーミングテストの手順（案）を策定 |
| 安全性評価の実施手法に関する検討 | 産学との意見交換<br>AI 安全性評価の運用に係る検討 | ７月目途に安全性の評価観点を整理 |
| 他国の関係機関との国際連携に関する業務 | 海外の関係機関との連携<br>付随する基礎調査 など | ３月に英米 AISI 所長や国内関係機関等との意見交換を実施。<br>５月までに今後の国際協力方針のセットを目指す。 |

（出所）AISI 関係府省庁等連絡会議（第 1 回）資料より丸紅経済研究所作成

（2）統一指針で目指すイノベーションと安全：AI 事業者ガイドライン

総務省と経済産業省は、2023 年 5 月 26 日に開催された AI 戦略会議で提示された AI に関する暫定的な論点整理を基に、2024 年 4 月に AI 事業者向けのガイドラインを公表した。このガイドラインは、AI ガバナンスに関する統一的な指針を示すことで、イノベーションの促進と AI ライフサイクル全体にわたるリスク緩和を両立する枠組みを構築することを目的としている。例えば、AI による意思決定・感情の操作等への留意、AI システム・サービス全般におけるプライバシーの保護、AI の判断にかかわる検証可能性の確保といった透明性について触れられている。このような指針は、AI 技術の急速な進展に伴い、社会や個人に与える影響が大きくなる中で、信頼性の高い AI システムの開発と利用を促進するために不可欠である。

ガイドラインの策定にあたり、広島 AI プロセスの議論やマルチステークホルダー・アプローチを重視し、総務省「AI ネットワーク社会推進会議」、経済産業省「AI 事業者ガイドライン検討会」、および関連するワーキンググループを活用して、産業界、アカデミア及び市民からの多様な意見を取り入れた。AI 事業者ガイドラインは、人間中心の AI 社会原則を基盤とし、過去の 3 つのガイドラインを統合し、国際的な動向を踏まえて策定された。また、AI ライフサイクルにおける具体的な役割を踏まえ、AI 開発者、AI 提供者、AI 利用者の 3 カテゴリーに分け、それぞれが遵守すべき指針や取組み事項が記載されている。

（3）イノベーションと権利保護：知的財産権の新たなフロンティア

AI のイノベーションを促進しつつ、創作者の権利を保護するために、内閣府は 2023 年 10 月 4 日に「AI 時代の知的財産権検討会」を設立した。この検討会の設立目的は、AI ツールの普及によって生じる生成 AI 関連の知的財産権をめぐる懸念やリスクに対処し、これらのツールの開発、提供、利用を促進することにより、日本の経済および社会の発展に貢献することである。2024 年 3 月 21 日に公表された「中間とりまとめの骨子（案）」では、著作権法以外の知的財産法との関係において、AI の学習段階では権利侵害に該当しないと考えられること、および生成・利用段階では知的財産権の侵害に該当するか判断することが示された。この内容は、6 月ごろに公表される政府の知的財産推進計画に反映される予定である。

著作権に関しては、環境変化に適応した著作物利用の円滑化を図り、新しいイノベーションを促進するため、2018 年 5 月 25 日に著作権法の一部を改正する法律が公布され、2019 年 1 月 1 日より、AI 学習・開発などのための著作物利用が原則として著作権者の許諾なく認められることとなった。しかし、生成 AI の飛躍的な進歩を背景に、AI の開発・利用による著作権侵害への懸念が高まっている。これを受けて、文化庁は文化審議会著作権分科会法制度小委員会での議論を重ね、2024 年 1 月 15 日に「AI と著作権に関する考え方について（素案）」を公表した。1 月 23 日から 2 月 12 日までパブリックコメントを受け付け、提出された 2 万 4,938 件の意見を踏まえ、3 月 15 日に修正版が了承された。ただ、この素案は法制度小委員会の見解を示すものであり、本考え方自体が法的な拘束力を有するものではない。

（4）自民党が検討する新たな法的枠組み：責任ある AI 推進基本法（仮称）

欧米で AI に関する法規制が進展していることを受け、自民党は AI の進化と実装に関するプロジェクトチームを設立した。このプロジェクトチームは、責任ある AI 利活用を推進するための法制度の

検討に取り組んでいる。この取り組みの一環として、ワーキンググループから「責任ある AI 推進基本法（仮称）」の素案が 2024 年 2 月 16 日に公表され、年内の法案提出が目指されている。

ワーキンググループの素案は、AI 技術の発展と社会への適用を促進するための基本的な方針や枠組みを定める。目的は、AI の利活用に伴う基本的人権をはじめとする国民の権利利益が侵害されるリスクを最小化し、AI によるイノベーションを含む健全な発展を通じて利益を最大化することである。これを実現するために、素案は安全で信頼性の高い AI の設計開発及び導入を促進し、人間中心の AI 利用を可能にする開かれた環境の整備を目指す。

特に、素案は特定の AI 基盤モデル開発者に焦点を当て、これらの開発者に対して体制整備や報告義務を定めている。具体的には、一定の規模や目的の AI 基盤モデル開発者を「特定 AI 基盤モデル開発者」として指定し、これらの指定された開発者に対しては、米国のボランタリーコミットメンツを参考にした体制整備の義務を課す。

**特定 AI 基盤モデル開発者の体制整備の義務**

| 米国のボランタリーコミットメンツ | 日本の責任ある AI 推進基本法 |
|---|---|
| セキュリティテストの実施 | 安全性検証を行う |
| リスクの管理に関する情報を共有 | リスク情報を共有 |
| サイバーセキュリティへの投資 | サイバーセキュリティへの投資 |
| 第三者による脆弱性の発見と報告 | 第三者による脆弱性等の検出と報告 |
| AI によって生成したコンテンツであることをユーザーが認識できるように取り組む | 生成 AI の利用を利用者に通知する仕組みの採用 |
| 能力、限界等の公開 | AI の能力、限界等の公表 |
| 社会的リスクに関する研究を優先 | 社会的リスクに関する研究推進 |
| 社会課題に対処するための開発、導入 | ――― |

（出所） 責任ある AI 推進基本法（仮称）の素案より丸紅経済研究所作成

素案では、特定 AI 基盤モデル開発者の体制整備義務を法律上は抽象的に定め、具体的な義務の内容は、民間事業者や業界団体が制定する行動規範によって詳細化されることを想定している。体制整備に関する報告義務については、特定 AI 基盤モデル開発者に定期的に体制整備義務の遵守状況を国または ASIS などの第三者機関に報告する義務を課す。国は民間やその他の関係者からの意見を収集でき、評価結果を公表し、必要に応じて是正措置を求めることができる。義務違反やインシデントが発生した場合には、報告徴求や立入検査をできる。また、報告義務や命令違反には、課徴金や刑罰を科すことも検討されている。さらに、民間規格の遵守が不十分な場合には、認証の取り消しや一時停止といった措置が講じられることも想定する。

## 5．まとめ

欧州で採択された AI 法は、AI システムがもたらすリスクの程度に基づき、規制の厳格性を決定するリスクベースのアプローチを採用している。このアプローチは、高リスク AI の使用を制限する一方で、イノベーションを阻害しないように意図されている。EU は、AI を人間中心の技術と位置づけ、

その最終目的を人間の幸福の向上に設定している。さらに、EU は健康、安全、および基本的人権の保護などの公共の利益を最優先し、これらを高い水準で保護することに重点を置いている。

一方、米国では、AI に関する包括的な連邦法の制定はまだない。安全、安心で信頼できる AI に関する大統領令に基づき、連邦政府が対応を行っている段階にある。米国は、技術的優位性の保護とイノベーションの促進に重点を置き、産業界との協力や自己規制を奨励している。

日本では、政府が AI 戦略を発表し、AI の安全で信頼できる活用が促進されるようガイドラインを策定する。AI に関する具体的な法律はまだ制定されておらず、開発力の向上と AI の多様なリスクへの対応強化に重点が置かれている。さらに、国際的な枠組みへの積極的な参加を通じて、世界的な AI ガバナンスの形成に貢献している。

現在の AI 関連の進展を考察すると、欧州と米国のアプローチにおけるシビル・ローとコモン・ローの違いが顕著に浮かび上がる。シビル・ローでは、成文法が主要な法源であり、裁判官は具体的な事件に対して成文法を適用する。この体系の利点は、法律が明文化されているため、法の予測可能性が高く、市民が法にアクセスしやすいことである。しかし、社会の急速な変化に法律が対応するのが遅れることが問題となる。一方、コモン・ローの根幹は判例主義にあり、歴史的な裁判所の判断が後続の類似事件に対する法の適用に重要な影響を及ぼす。この長所は、実際の事件を基に法が進化することで、社会の変遷に柔軟に対応できる点にある。しかし、判例の複雑な相互作用により、法の透明性が損なわれることもある。

欧州ではシビル・ローの枠組みの下で AI 法が導入され、法の予測可能性を向上させつつも、技術革新への迅速な対応が課題になり得ると指摘されている。米国では、コモン・ローの柔軟性が AI 技術の急速な進化に対応できる可能性があるが、法の透明性と予測可能性の問題は生じ得る。明治時代にシビル・ローを採用した日本は、AI については米国に倣った政策をとりつつも、「広島 AI プロセス」という取り組みを通じて、国際的枠組みの構築に貢献している。

技術革新の速度と社会的影響の大きさを考えると、国際的な枠組みの整備と相互協力の促進は、今後さらに重要になると予想される。AI に関する規制の制定と施行が進むにつれ、これらの変化がビジネス環境に与える影響はさらに増大するであろう。グローバルな市場で事業を展開する企業にとって、国際的なトレンドに柔軟に対応することが不可欠である。技術革新と社会的課題の間でバランスを取りながら推進される各国の政策と国際協力に関しては、常に最新の動向や法律、ガイドラインを確認しておくことが求められる。

以　上

**丸紅経済研究所**

〒100-8088　東京都千代田区大手町一丁目４番２号
https://www.marubeni.com/jp/research/