IISE 調査研究レポート（No.1）

「欧州における AI 規制の動き～日本に必要な AI 規制とは～」

（2023 年度「プライバシー/DFFT を巡る社会課題と政策動向に関する調査研究」報告書より抜粋）

2024 年 4 月
国際社会経済研究所　主幹研究員　小泉 雄介

EU で AI 法（AI 規則）の制定に向けた作業が進められ、米国でも AI 大統領令が発令される中、日本における AI に対する法規制については、2023 年内の G7 広島 AI プロセスの進行中は、日本政府や有識者は技術進歩の急速な AI／生成 AI に対する法規制に消極的であり、事業者が自主的に遵守する新たな AI 事業者ガイドライン（ソフトロー）での対応を志向していた。また、既存法令での対応は既に一部でなされており、個人情報保護委員会が 2023 年 6 月に OpenAI に対して個人情報保護法に基づく注意喚起[1]を行った（ChatGPT の利用者から本人同意なく要配慮個人情報を取得しないように求めるもの）。また、文化庁は 2024 年 1 月～2 月に「AI と著作権に関する考え方について（素案）」に関するパブコメ募集を行い、3 月には正式版[2]を公表している。さらに、AI リスクに対してその他の既存法令（業法など）や技術で対応できない部分があるか否かを明確化するための作業も各省で行われている。

G7 広島 AI プロセスと並んで、日本政府による取組みとして注目された AI 事業者ガイドラインに関しては、「AI がもたらす社会的リスクの低減を図る（・・・）ために、関係者による自主的な取組を促」すものとされているが、もし真面目な事業者のみが本ガイドラインを愚直に遵守し、その他の事業者が遵守せず、その結果として両者の事業効率に差が生じたり（例えば真面目な事業者の AI 開発速度が遅くなったり）、両者のリスク対策に差が生じたり（例えばガイドラインを顧慮しない事業者から重大な侵害事案が発生）した場合、AI 事業者ガイドラインの発行意義が問われることになる。

日本社会が（主に外国製品に起因する）AI リスクを丸被りせず、またガイドラインを真面目に遵守する事業者が不利な立場にならないようにするためには、①AI 事業者ガイドライン（の一部）に対する何らかの履行確保の手段、もしくは②ガイドラインを遵守する事業者に対するインセンティブを整備することが、今後の日本政府に求められる施策となるだろう。

具体的には、①高度な AI システムの開発者を対象とする新たな AI 規制法の制定、②日本政府が調達する AI システムの入札条件となる AI ガバナンス認証制度[3]の創設など、いくつかのオプションが考えられる。

欧州委員会が AI 法案を提案した理由は「AI ホワイトペーパー（White Paper on Artificial

[1] https://www.ppc.go.jp/news/careful_information/230602_AI_utilize_alert/。
[2] https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_07/pdf/94024201_01.pdf。
[3] 認証基準は AI 事業者ガイドラインに準じたものとし、政府が認定した認証機関が第三者認証を行う。また、外国との相互承認のためには政府の関与（法のフック）が必要である。

Intelligence: a European approach to excellence and trust)」[4]に記載されているが、そのうち大きなものは、「AI がもたらすリスクの中には既存法令でカバーできないものがある」「AI システムは市場投入後、学習等により機能が変化しうるため、市場投入時には顕在化していなかった新たな安全性リスクを生み出す恐れがある」という理由である。これらについては日本の社会環境にもそのまま当てはまるため、日本においても①の新たな AI 規制法を立法する意義は十分にあるものと思われる。

また、高度な AI システム（大規模な基盤モデル等）を開発する事業者の多くが米国 AI 事業者であり、国内の AI 提供者・利用者が多くの場合、それら米国事業者の提供する AI システム（基盤モデル等）を利用せざるを得ず、米国などで顕在化している AI リスクの事例が国内でも発生しうることを考慮すると、大規模な基盤モデル等を開発する外国 AI 事業者を主な対象として日本の国内法で規制することについては、日本国民や事業者の一定の理解が得られるものと考えられる。ただし欧米と異なり、日本においては AI のリスクが顕在化した事案や事件、犯罪などがほとんど発生していない。AI 事業者ガイドライン案の別添にリスト化されている「AI によるリスク」の事例もほとんどが欧米における事例[5]であり、日本の事例はリクナビの内定辞退率事案ぐらいである。しかもリクナビ事案については既存法令（個人情報保護法と職業安定法）で対処がなされている。今後、日本政府が①（高度な AI システムの開発者を対象とする新たな AI 規制法の制定）のオプションを取る場合には、日本において AI リスクが顕在化した事案・事件・犯罪がほとんどなく[6]、「立法事実」が乏しいため、立法に当たっての正当性をどう根拠付けるかは一つの課題となりうる。

①の高度な AI システムの開発者に対する法規制内容としては、米国の自主ルール（Voluntary Commitments）を参考に、自民党 AI の進化と実装に関するプロジェクトチーム（AIPT）が 7 項目の体制整備を課す私案「責任ある AI 推進基本法（仮）」[7]を提言している。米国 AI 事業者に追従することになる嫌いはあるものの、G7 広島 AI プロセスの国際指針・国際行動規範も米国の自主ルールを参照していることもあり、大きな流れとしては米国自主ルールに基づかざるを得ない状況にあると言える。

②（日本政府が調達する AI システムの入札条件となる AI ガバナンス認証制度）のオプションについては、必ずしも米国自主ルールに準拠する必要はない。国際規格である ISO/IEC 42001（AI

---

[4] https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en。

[5] Amazon の人事採用バイアス、ケンブリッジアナリティカ事件、Tesla の自動運転車死亡事故、Microsoft のチャットボット Tay のヘイトスピーチ、Apple のクレジットカード利用限度額の性差別、米国アリゾナ州での娘の声を AI 合成した身代金要求事件、ChatGPT のハルシネーション（詐欺と横領の容疑）に対する米国ジョージア州男性の名誉棄損訴訟、米国ニューヨーク州弁護士の ChatGPT 利用による偽判例の引用、米国の「国防総省付近で爆発が起きた」とする生成 AI で作られた偽画像の流布、米国アーティストによる画像生成 AI に対する著作権侵害の集団訴訟など。

[6] 数少ない事案としては、三重県の児童相談所が児童保護の判断に AI の評価を使い「保護率 39%」だったので積極的対応をしなかったところ、2023 年 5 月に 4 歳児童が虐待で死亡してしまった事件がある。https://www3.nhk.or.jp/news/html/20230711/k10014125711000.html。

[7] https://note.com/akihisa_shiozaki/n/n4c126c27fd3d。

マネジメントシステム）は、AI 法の EU 整合規格にも反映される見込みであり、AI 事業者ガイドラインにおいても参考にされている[8]。この「政府調達 AI の入札条件としての認証制度」については、マネジメントシステムとしては ISO 規格、個々の AI システムの開発時・提供時の配慮事項としては AI 事業者ガイドラインに準拠する案などが考えられる。

表　EU の AI 法に挙げられた 3 つの AI リスクと欧米日の対応（筆者作成）

AI のもたらすリスク		EU AI 法	米国 AI 大統領令	日本 自民党 AIPT 私案
従来からの AI 一般のリスク	(1)身体等の安全や健康に対するリスク	(1)既存の製品規制法令の修正で対応（EU 整合規格への適合性評価）	— （州法等が存在？）	— （業法改正等で今後対応）
	(2)基本的人権に対するリスク	(2)禁止 AI またはハイリスク AI としてプロバイダー等を規制（EU 整合規格への適合性評価）	— （州法・条例で個別分野を規制：人事採用、顔識別等）	— （既存法令で一部対応：リクナビ事案等） （政府調達 AI の入札条件として認証制度も想定しうる）
(3)生成 AI・基盤モデルのリスク	開発者に起因するリスク（公衆衛生・安全・治安・基本的人権・社会全体に対するリスク）→シンギュラリティ／AGI も視野？	(3)GPAI モデルやシステミックリスクのある GPAI モデルのプロバイダーを規制（EU 整合規格または実践規範）	開発者を規制（NIST がレッドチームテストの規格を策定）	開発者を規制（規格や行動規範は民間が決める共同規制）
	利用者に起因するリスク	AI 生成コンテンツの開示義務のみ （デジタルサービス法で偽情報等の流通を規制）	— （通信品位法の改正法案が提出）	— （プロバイダー責任制限法改正で偽情報対策強化）

なお EU の AI 法では、AI がもたらすリスクとして、大きく 3 つのリスクを想定している（上表参照）。そのうち 2 つは従来から捉えられてきた AI リスク（1）（2）であり、もう 1 つは 2023 年にクローズアップされた生成 AI・基盤モデルがもたらすリスク（3）である。従来から捉えられてきた AI リスクのうち「(1) 身体等の安全や健康に対するリスク」に対しては、産業機械や医療機器など既存の EU 製品規制法令の（整合規格の）修正で対応し、「(2) 基本的人権に対するリスク」に対しては、禁止 AI やハイリスク AI（スタンドアロン AI）を指定・分類して禁止もしくは新たな整合規格でプロバイダーやディプロイヤーを規制しようとしている。

AI 法では「(3) 生成 AI・基盤モデルがもたらすリスク」に対しては、開発者に起因するリスク

[8] 2023 年 12 月 18 日には国際規格として、ISO/IEC 42001: 2003（AI に関するマネジメントシステム規格）が発行された。この ISO/IEC 42001 は AI を利用した製品やサービスを提供する組織が責任を持ってその役割を果たすための要求事項を示すものである。EU の AI 法の（ハイリスク AI の）適合性評価に用いられる整合規格（Harmonised standards）については、欧州委員会から欧州の標準化団体（CEN と CENELEC）に対して策定が委託されているが、この AI 法の整合規格についても ISO 規格が反映されるものと考えられる（https://webdesk.jsa.or.jp/common/W10K0620?id=1101）。したがって、「日本政府が調達する AI システムの入札条件となる AI ガバナンス認証制度」において、この ISO/IEC 42001 を準拠規格とすることは、EU の AI 法への対応を視野に入れる上でも、一つの選択肢となりうるだろう。

（公衆衛生・安全・治安・基本的人権・社会全体に対するリスク）について、GPAI（汎用目的 AI）モデルやシステミックリスクのある GPAI モデルのプロバイダーを整合規格（整合規格ができるまでは実践規範）で規制しようとしている。なお、生成 AI・基盤モデルの利用者（ディプロイヤー）に起因するリスクに対しては、AI 法の中では（AI 生成コンテンツの開示義務以外には）規制されておらず、デジタルサービス法（DSA）で偽情報等の流通が規制されている。

米国の AI 大統領令や日本の自民党 AIPT 私案「責任ある AI 推進基本法」については、上記（1）～（3）のリスクのうち、「（3）生成 AI・基盤モデルがもたらすリスク」の開発者に起因するリスクに対する規制しか盛り込まれていない（上表参照）。日本では、「（1）身体等の安全や健康に対するリスク」については自動車や医療機器などの既存の業法の改正で対応するものと思われるが、バイアス・自動意思決定・人間の関与・透明性など「（2）基本的人権に対するリスク」に関しては、現状では AI 事業者ガイドラインに基づく事業者の自主的な取組に委ねられることとなる。（2）については今後、前述のように、（政府調達 AI の入札条件として）AI 事業者ガイドライン／ISO に基づく認証制度を整備することにより、自主的な取組を促進することが求められるのではないだろうか。

以　上