論説

「AI 規制論」のコペルニクス的転回
――現代の一般的規制モデルの構築に向けて――

弁護士・京都大学特任教授　東京大学法科大学院客員准教授
羽深宏樹

Ⅰ．はじめに

2012 年にディープラーニングが飛躍的発展を遂げてから数年のうちに，AI は社会の様々なサービスに実装されるようになり，それに伴って多方面で AI によるリスクが顕在化してきた。自動運転車による死亡事故，金融取引アルゴリズムによる相場の急変動，採用アルゴリズムによる性・人種差別，生成 AI で作成したディープフェイク画像・動画の拡散など，その例を挙げればきりがない。こうした様々なリスクを受けて，AI やこれを組み込んだシステムがユーザーや第三者にもたらすリスクに関する法規制（以下，これを「AI 規制」と呼ぶ[1)]）に関する議論が国

1)　本稿では，AI の開発や利用に伴う安全規制を中心に扱う。そのため，取引法（民法，商法等），知的財産法（著作権法，特許法等），データ法（個人情報保護法，不正競争防止法等），競争法（独占禁止法，不正競争防止法等），デジタルプラットフォーム関連法（デジタルプラットフォーム取引透明化法，スマートフォンソフトウェア競争

内外で急速に進んでいる。日本では，AI に関する制度全般の方針を示すものとして，AI 戦略会議によって 2024 年 5 月に AI を巡る制度設計の方向性が整理され[2]，2025 年 2 月には AI 制度研究会によって包括的な中間とりまとめが公表された。そのほか，自動運転[3]，医療機器[4]，知的財産権[5] などの分野においても，政府主導で法整備や法解釈の整理が行われている。国外では，EU において AI を包括的に規制する AI 法が 2024 年 8 月に発効し[6]，同年 9 月には，米国カリフォルニア州において生成 AI に関する複数の法律が成立したほか[7]，英国でも 5 月に「自動運転車法」[8] が制定されるなど[9]，盛んに立法が行われている。また，国際的な合意形成も進められており，その中には，2023 年の広島 G7 サミットにおける「広島 AI プロセス」の立ち上げ[10]，欧州のみならず米国やカナダ，そして日本も参加する「AI 条約」の調印[11]，経済協力開発機構（OECD）による「AI 原則」[12]（以下，「OECD AI 原則」）の更新など，様々な枠組が構築されている[13]。

これらの目まぐるしい動きは，社会に「AI」という特別な技術が入ってくるために，その対応に必要な範囲で法改正や合意形成を進めているものだ，と思われがちである。しかし，本稿では，それとは全く逆の見方を提唱する。すなわち，現在 AI 規制として議論されていることは，AI という技術に固有の問題ではなく，既存の規制に内在していた問題を，AI が浮き彫りにしたに過ぎないという見方だ。したがって，我々が最終的に目指すべきなのは，AI に関する新法を制定したり，

促進法等），環境法，安全保障等に関する具体的な論点ついては，AI に関する法規として非常に重要ではあるものの，本稿の中では必要に応じて簡単に触れる程度にとどめる。

2) 内閣府 AI 戦略会議「『AI 制度に関する考え方』について」（2024 年 5 月）。

3) デジタル庁 AI 時代における自動運転車の社会的ルールの在り方検討サブワーキンググループ「AI 時代における自動運転車の社会的ルールの在り方検討サブワーキンググループ報告書」（2024 年 5 月）（https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/1fd724f2-4206-4998-a4c0-60395fd0fa95/9979bca8/20240523_meeting_mobility-subworking-group_outline_04%20.pdf，2024 年 10 月 5 日最終閲覧）。

4) 厚生労働省「令和 4 年度『プログラム医療機器の特性を踏まえた薬事承認制度の運用改善検討事業』報告書　プログラム医療機器の特性を踏まえた適切かつ迅速な承認及び開発のためのガイダンス」（2023 年 3 月）（https://www.mhlw.go.jp/web/t_doc?dataId=00tc7714&dataType=1&pageNo=1，2024 年 10 月 5 日最終閲覧）。

5) 知的財産戦略本部「AI 時代の知的財産権検討会中間とりまとめ」（2024 年 5 月）（https://www.kantei.go.jp/jp/singi/titeki2/chitekizaisan2024/0528_ai.pdf，2024 年 10 月 5 日最終閲覧）。

6) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX-%3A32024R1689)

7) JETRO「米カリフォルニア州知事，ディープフェイク関連の 9 つの AI 法案に署名」(https://www.jetro.go.jp/biznews/2024/09/a5ec5ef37532a9d9.html)。なお，同時期に審議されていた，大規模基盤モデルを対象とする「最先端人工知能（AI）システムのための安全でセキュアな技術革新法（SB1047）」（カリフォルニア AI 安全法）については，州知事が拒否権を行使したために廃案となった。

8) Automated Vehicles Act 2024, https://www.legislation.gov.uk/ukpga/2024/10/pdfs/ukpga_20240010_en.pdf, last visited Oct 5, 2024.

9) SB-1047 Safe and Secure Innovation for Frontier Artificial Intelligence Models Act, https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB1047, last visited Oct 5, 2024.

10) https://www.soumu.go.jp/hiroshimaaiprocess/documents.html, 2024 年 10 月 5 日最終閲覧

11) Council of Europe, Council of Europe opens first ever global treaty on AI for signature, Sep. 5, 2024, https://www.coe.int/en/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature, last visited Oct 5, 2024, Canada and Japan sign Council of Europe's first ever global treaty on AI, February 11, 2025, https://www.coe.int/en/web/portal/-/canada-and-japan-sign-council-of-europe-s-first-ever-global-treaty-on-ai, last visited Feb 14, 2025.

12) OECD, Recommendation of the Council on Artificial Intelligence" ,amended in 2024, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449,.

13) 以上を含む世界の動きについては，拙稿参照。Hiroki Habuka et al. , *Shaping Global AI Governance: Enhancements and Next Steps for the G7 Hiroshima AI Process*, May 24, 2024, Center for Strategic & International Studies, https://www.csis.org/analysis/shaping-global-ai-governance-enhancements-and-next-steps-g7-hiroshima-ai-process, last visited Oct 5, 2024.

法令の中にAIに関する条項を追加したりすることではなく，規制一般の内容や運用を，AIの使用の有無を問わず全般的に見直すことではないか，というのが本稿の主張である。そして，現在「AI規制論」として語られていることは，そうした法制度の一般的な見直しにあたって重要な手がかりとなる。

以上を導くため，本稿では次のように議論を進める。まず，昨今「AI規制論」として議論される，一般的なAIのリスクの内容やそれに対する規制のアプローチを整理する（Ⅱ）。次に，AIと人間のリスクが相対的かつ不可分であることを説明し，その帰結として，AI規制の文脈で課題とされるリスクや規制のアプローチが，AI以外の文脈で既に論じられてきた内容と概ね共通していることを示す（Ⅲ）。続けて，AI規制論こそが，不確実性を増す現代社会において法の支配を実質化するための新たなアプローチであることを示し，そのためこれをAIという特定の技術分野に限定するのではなく，一般的な規制論に昇華させるべきであると主張する（Ⅳ）。そして，AI規制論を梃子としてあらゆる規制をアップデートしていくことに関する意義と課題を述べる（Ⅴ）。最後に，全体の議論をまとめる（Ⅵ）。

なお，本稿における事実に関する記述は，脱稿時点である2025年2月における最新情報である。AI規制の分野では瞬く間に状況がアップデートされるため，最新の法令動向などについては，各国の政府ウェブサイトや信頼できるメディア等を参照されたい。

Ⅱ.「AI規制論」の概要

本節では，過去数年間にわたって展開されてきた，AI規制に関する国内外の一般的な議論（以下，「AI規制論」という）の概要を説明する。

1 AIリスク

ある行為やシステムに関する規制を議論する際には，まず，立法目的を裏付ける一般的事実（立法事実）の検討が必要になる[14]。すなわち，その規制がどのようなリスクに対処し，何を達成することを意図したものかを明確にしなければならない。

AIが社会に及ぼすリスクについては，2010年代の後半頃から，国内外で様々な議論が展開されてきた。

たとえば，冒頭で触れた内閣府AI戦略会議による「『AI制度に関する考え方』について」では，制度的な検討を要するリスクとして，(i) 製品・サービスの安全性，(ii) 人権侵害（プライバシーや公平性など），(iii) 安全保障・犯罪増加，(iv) 財産権の侵害，(v) 知的財産権の侵害，(vi) その他（雇用への影響，暴走，利益の集中，多様性），というリスクが挙げられている。

また，AIを包括的に規制するEUのAI法は，その目的として，(i) 健康，(ii) 安全，(iii) 民主主義，(iv) 法の支配，(v) 環境の保護を掲げている[15]。また，身体的，心理的，社会的，経済的な害に関する言及もある[16]。同法では，こうしたリスクのレベルに応じて，禁止されるAIの利用，ハイリスクAIシステム，透明性を求められるAIシステム，といったカテゴリーごとの規制がなされている。

国連のAI諮問機関（AI Advisory Body）が指摘するAIリスクは，さらに多岐にわたる。そこでは，上述のようなリスクに加えて，安全保障（兵器目的でのAI利用）やグループの孤立，多様性や社会の結束への影響にも言及されており，その上で，あらゆる時代におけるAIリスクの包括的なリストを作成するのは不可能であると宣言している[17]。

14) 芦部信喜 著（高橋和之補訂）『憲法（第8版）』409頁（岩波書店，2023）。
15) EU AI法前文（1）。
16) EU AI法前文（5）。
17) United Nations AI Advisory Body, *Interim Report: "Governing AI for Humanity"*, Dec. 2023, 8-9 (https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf), last visited Oct 5, 2024.

このように，AI がもたらすリスクは，個人レベル（生命・健康，プライバシー，有形・無形の財産権，自律性，雇用，その他の基本的人権等），コミュニティレベル（差別，多様性等），国家レベル（安全保障，民主主義，法の支配等），そしてグローバルレベル（気候変動，力の集中と地域格差等）に至るまで，極めて幅広い。

2　AI 原則

こうした様々な AI のリスクを踏まえ，2016 年頃から，世界各国の政府や国際機関，非営利団体や企業などが続々と，AI に関するポリシーや原則を公表した。これらのポリシーや原則は，政府によるものだけでも世界で 1000 以上存在する[18]。そこでは概ね，人権や民主主義などの基本的価値や，公平性，プライバシー，安全性，セキュリティ，透明性，アカウンタビリティといった原則の重要性が強調された。上述の OECD AI 原則（2019 年）や，日本政府の「人間中心の AI 社会原則」（2019 年）[19]，国連教育科学文化機関（UNESCO）の「AI 倫理に関する勧告」（2021 年）[20] などが例として挙げられる。

ハーバード大学の Berkman Klein Center の調査によれば，2019 年までに世界で公表された 40 近くの AI 原則の内容は，(i) プライバシー，(ii) 安全性とセキュリティ，(iii) 透明性・説明可能性，(iv) 公平性・無差別，(v) アカウンタビリティ，(vi) 人間による技術のコントロール，(vii) 専門家の責任，(viii) 人間的価値の促進に分類できるという[21]。以下では，これらを総称して「AI 原則」という。

3　規範形成の状況

このような AI 原則は非常に抽象的な理念であるため，これらを実際の AI システムに落とし込むためには具体的な規範が必要となる。国内外では，ハードローとソフトローの両面で整備が進められている。ハードローとは法的拘束力を有し，違反すると制裁が科されるルールであるのに対し，ソフトローとはそのような法的拘束力がなく，従うかどうかは事業者の自主的な判断に委ねられる規範をいう。

ハードローとして，現時点で最も包括的に AI を規制しているのは，EU の AI 法であろう。2024 年に成立した AI 法は，事業分野を問わずに適用され，禁止される AI の利用，ハイリスク AI システム，透明性を求められる AI システム，といったカテゴリーごとに AI システムの開発者や利用者等に様々な義務を課している[22]。日本には，EU のような包括的な規制は存在しないが，分野別には，自動運転システム（道路運送車両法）[23] やプログラム医療機器（薬機法）[24] などの個別法によって法的拘束力のある規制が及んで

18)　OECD *National AI policies & strategies* (https://oecd.ai/en/dashboards/overview), last visited Oct 16, 2024.

19)　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）（https://www8.cao.go.jp/cstp/aigensoku.pdf，2024 年 10 月 5 日最終閲覧）。

20)　UNESCO, *Recommendation on the Ethics of Artificial Intelligence*2022, https://unesdoc.unesco.org/ark:/48223/pf0000381137, last visited Oct 5, 2024.

21)　Fjeld, Jessicaet al. , *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI*, Feb. 14, 2020, 2020-1 BERKMAN KLEIN CENTER RESEARCH PUBLICATION , *available at* https://ssrn.com/abstract=3518482, last visited Oct 5, 2024 or http://dx.doi.org/10.2139/ssrn.3518482, last visited Oct 5, 2024.

22)　EU AI 法の詳細については，古川直裕・羽深宏樹ほか「EU AI 法概説―(1)～(9・完)」NBL1269 号（2024 年 7 月）～ 1271 号（2024 年 8 月），1273 号（2024 年 9 月）～ 1278 号（2024 年 11 月）連載を参照。

23)　道路運送車両の保安基準 48 条，道路運送車両の保安基準の細目を定める告示 150 条の 2。

24)　医師の診察を手助けする AI などのシステムは「プログラム医療機器（SaMD）」に該当するため，薬機法が定める承認手続が必要となる。そのためには，治験や審査に 5 年以上かかることもあるが，これを開発から 1 年以内に販売できるようにする制度改正が検討されている。その際，AI が，現場で使われた後も随時アップデートされていくことを踏まえ，まずは最低限の審査でいったん販売を許可し，使用後のデータから再度承認できるかどうかを判断する 2 段階の制度が検討されている。厚生労働省「令和 4 年度『プログラム医療機器の特性を踏まえた薬事承認制度の運用改善検討事業』 報告書プログラム医療機器の特性を踏まえた適切かつ迅速な承認及び開

いる。2024年には，英国で「自動運転車法」が，米国カリフォルニア州では生成AIの学習データの透明化に関する法律が成立するなど，世界中でAIに対する規制の動きが進んでいる。

なお，この点について，EUがハードローのアプローチを採っているのに対し，日本ではソフトローのアプローチを採っている，といった説明がなされる場合もあるが[25]，そのような整理はミスリーディングといえる。日本のみならず，諸外国においても，自動車や医療分野を中心に，セクターごとにAIに関するハードローが置かれている場合が多い。一方で，「ハードローアプローチ」の典型とされるEUのAI法においても，その具体的な履行態様については多くが標準（ソフトロー）に委ねられており，ハードローが規定しているのは，あくまでも規律の大枠である。

ソフトローについても各国で整備が進められている。日本では，2024年に「AI事業者ガイドライン」が公表された。国際標準化機構（ISO）も，2020年以降，次々とAIのガバナンスに関する国際規格を発行しており，2023年12月にはAIのマネジメントシステム規格であるISO/IEC42001が公表された。米国の国立標準技術研究所（NIST）は，2023年に「AIリスクマネジメントフレームワーク」[26]を公開し，2024年には，その枠組みを生成AIに適用した「AIリスクマネジメントフレームワーク：生成AIプロファイル」[27]を公表した。EUでも，現在，AI法に対応するためのEU整合規格の整備が進められている[28]。

4　AI規制のアプローチ

このように，AI規制の取組は各国によって異なる。より具体的な差異と共通点については別稿[29]を参照されたいが，法整備に関しては，概ね以下のような共通項を示すことができる。（以下，これらを総称して「AI規制のアプローチ」という。）

①　リスクベースアプローチ

AIのリスクは多岐にわたるため，全てのAIに適用されるような単一の解決策（one-size-fits-all solution）を適用することは難しい。そのため，AIがもたらすリスクの高さに応じて異なる対応を行う「リスクベースアプローチ」が必要であるという点において，各国の立場は一致している。

もっとも，この「リスクベースアプローチ」という言葉は，異なる文脈で用いられることに注意が必要だ。もともとリスクベースアプローチという用語が用いられていた金融規制の分野では，規制当局のリソースを最も高いリスクに優先的に割り当ててリスクに対応することを意味するものとして用いられてきた[30]。

これに対し，EUのAI法では，立法段階において「リスクベース」という用語が用いられている。すなわち，AIのリスクに応じて「禁止されるAIの利用」，「ハイリスクAIシステム」，「透明性を求められるAIシステ

発のためのガイダンス」（2023年）（https://www.mhlw.go.jp/web/t_doc?dataId=00tc7714&dataType=1&pageNo=1, 2024年10月5日最終閲覧）。

25) たとえば，内閣府AI戦略会議が2024年5月に公表した前掲注2)「『AI制度に関する考え方』について」では，「（AI規制について）我が国はソフトローによる対応を中心に行ってきた。一方，刑法，個人情報保護法などAIか否かに関わらず適用される法律はあるが，欧米のような大規模あるいはリスクの高いAI等に適用されるハードローは存在しない。」という記述がある。

26) NIST, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, Jan. 2023, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf, last visited Oct 5, 2024.

27) NIST, *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*, Jul. 2024, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf, last visited Oct 5, 2024.

28) European Commission, "*C(2023)3215 – Standardisation request M/593*" (https://ec.europa.eu/growth/tools-databases/enorm/mandate/593_en), last visited Oct 16, 2024.

29) 前掲注13

30) Julia Black & Robert Baldwin, *Really responsive risk-based regulation*, 32(2) LAW AND POLICY, at1 (spring 2010).

ム」，といったカテゴリーごとの規制がなされている点をもって「リスクベース」と称している[31]。ただし，これは立法手段における比例原則[32]の意味に近いであろう。すなわち，同法では，AIがもたらす多様なリスクについて，適合性・必要性・狭義の比例性を検討した結果として，複数のリスクカテゴリーごとに異なる規範が適用されることになったものと理解できる。そうだとすれば，自動車，医療，金融といった分野ごとに異なる規制を置いている日本，米国，英国その他の国においても，各リスクに応じて適用される法規範が異なるという意味で，立法段階の「リスクベースアプローチ」が採られているといえよう。

さらに，企業の内部統制の文脈で「リスクベース」という言葉が用いられることもある。たとえば，日本のAI事業者ガイドラインでは，各事業者が「予め事前に当該利用分野における利用形態に伴って生じうるリスクの大きさ（危害の大きさ及びその蓋然性）を把握したうえで，その対策の程度をリスクの大きさに対応させる」ことを「リスクベースアプローチ」と呼んでいる[33]。

このように，「リスクベースアプローチ」という言葉は，監督，立法，そして内部統制といった複数の文脈で使用されているが，これらは，どれかを選択的に適用するのではなく，重層的に適用することで，社会全体としてリスクの大きな分野に重点的にリスク対応がなされることを目指すものであろう。たとえば，AIガバナンスに関する重要な国際合意文書である「G7広島AIサミット包括政策枠組」[34]では，政策形成と企業の内部統制レベルの双方において，「リスクベースアプローチ」という用語が用いられている[35]。

リスクベースアプローチは，その考え方について異論を挟む余地はないと思われるが，難しいのはその実装である。すなわち，異なる種類のリスクをどう比較するのか，予見困難なリスクの確率をどのように算定（仮定）するのか，また日々更新されるリスク状況をどのように迅速に政策や内部統制に活かすのか，といった点は，実務上きわめて困難な検討課題である。

② ハードローによる大枠の設定とソフトローによる具体化

変化が速く複雑なAIの世界において，規制対象となるAIに関する技術的な要素を書き込むことは現実的ではない。そこで，法律では，システムや事業者が満たすべき要件の骨子（リスクマネジメント，データガバナンス，サイバーセキュリティ，記録，報告等）や，市場参入要件を満たすためのプロセス（適合性評価の自己宣言や第三者認証，規制当局による免許等）についてのみ定めておき，実際の判断基準は，法的拘束力のない標準やガイドライン，行動規範（Code of Conduct）などのソフトローで設定することが一般的である。たとえば，EUのAI法において，AIのリスクマネジメントシステムやデータガバナンス，記録管理等に関する法文は簡潔であり[36]，その下で詳細な標準の開発が進められている[37]。廃案となったカリフォ

31) European Commission, *Artificial Intelligence – Questions and Answers* (https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683), last visited Oct 16, 2024.

32) 立法手段における比例原則とは，①規制手段が規制目的と合理的に関連すること（適合性），②規制手段が規制目的の達成にとって必要最小限であること（必要性），及び③規制手段の投入によって得られる利益と失われる利益のバランスが均衡を保っていること（狭義の比例制）をいう。（渡辺康行ほか『憲法Ⅰ――――基本権（第2版）』78頁〔松本和彦〕（日本評論社，2023））。

33) 総務省＝経産省「AI事業者ガイドライン」3頁（2024年4月19日）（https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240419_1.pdf，2024年10月5日最終閲覧）。

34) Hiroshima AI Process G7 Digital & Tech Ministers' Statement, at 1-2 Hiroshima AI Process Comprehensive Policy Framework, Dec. 1, 2023, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document02_en.pdf, last visited Oct 5, 2024.

35) 前者については本文のパラグラフ3に，後者については別添1“Hiroshima Process International Guiding Principles for All AI Actors”2.Vにおいて言及がある。

36) EU AI法9，10，12条など参照。

37) European Commission, Draft standardisation request to the European Standardisation Organisations in sup-

ルニア AI 安全法でも，AI 開発者が，各種の標準を踏まえたベストプラクティスとガイダンスを考慮することを求めていた。その他，AI に関する国際規格の現状と今後の課題については，米国 NIST による取りまとめ[38]が参考になる。

これらの詳細をソフトローに委ねることの意義は，ハードローよりも迅速に更新が可能であるため状況変化に対応しやすいことだけではない。ソフトローには法的拘束力がないため，そこで示された手順に従わなくても，ハードローの定める目的さえ達成できていれば，法令違反になることがない。その意味で，事業者による革新的なコンプライアンス手法を許容できるというメリットがある。但し，実際にはソフトローに従う限り法的または事実上の合法推定が働く場合もあるため[39]，ソフトローが実務的にハードローに近い機能をもつこともある。

③ プロダクト要件とマネジメント要件の組み合わせ

AI 規制には，AI システム自体が備えるべき性能や構造に関する要件（プロダクト要件）と，AI システムを開発・運用する事業者が遵守すべき要件（マネジメント要件）という 2 つの系統の義務が課されることが多い。

プロダクト要件は，いわゆる製品安全を扱うものだ。たとえば，日本では，疾病の診断・治療等を目的としたソフトウェア（医療機器プログラム）について，そのリスクに応じて大臣承認または第三者認証を必要としている[40]。また，自動運転車に搭載される自動運行装置についても，型式指定が行われる[41]。

EU では，AI 法を，製品安全に関する一連の政策パッケージである NLF（New Legislative Framework）の一環に位置付けている。NLF には，AI 法のほかに，機械やバッテリー，医療機器規則等の製品の安全規制が含まれている[42]。AI 法では，ハイリスク AI システムに関する要件が 8 条から 15 条にかけて定められており，そこでは，リスクマネジメントシステム，データガバナンス，技術文書，記録保持，透明性，人間による監視，正確性とセキュリティなどが挙げられている。

他方，マネジメント要件とは，AI システムそれ自体の要件というより，AI システムを開発したり運用したりする事業者に課せられる義務だ。たとえば，薬機法では，医療機器等の製造業者に対して，製造管理及び品質管理の基準の遵守[43]や，製造及び試験に関する記録の保管[44]が義務付けられている。EU の AI 法では，ハイリスク AI システムのプロバイダに対して，品質マネジメントシステムの確立が義務付けられるほか，技術文書や品質マネジメントシステム関連文書，自動生成ログの保管などが義務付けられている[45]。

法的義務ではないが，AI マネジメントシステムについては，2023 年に，上述の ISO/IEC 42001[46]という国際規格が発行された。これは，世界的に多くの企業が認証を取得している ISO/IEC 27001（情報セキュリティマ

port of safe and trustworthy artificial intelligence, Dec. 5, 2022, https://ec.europa.eu/docsroom/documents/52376, last visited Oct 5, 2024.

38) NIST AI 100-5, *A Plan for Global Engagement on AI Standards*, Jul. 2024, https://doi.org/10.6028/NIST.AI.100-5, last visited Oct 5, 2024.

39) EU AI 法 40 条。

40) 薬機法 23 条の 2 の 5，同 23 条の 2 の 23。

41) 道路運送車両法 41 条 1 項 20 号，75 条の 3 第 1 項。

42) European Commission, New legislative framework, https://single-market-economy.ec.europa.eu/single-market/goods/new-legislative-framework_en, last visited Oct 5, 2024.

43) 薬機法施行規則第 114 条の 58，医療機器及び体外診断用医薬品の製造管理及び品質管理の基準に関する省令（平成 16 年厚生労働省令第 169 号）（以下「QMS 省令」）83 条。

44) 薬機法施行規則第 114 条の 51，QMS 省令 9 条，68 条。

45) EU AI 法 17 条ないし 19 条。

46) ISO, ISO/IEC 42001:2023, https://www.iso.org/standard/81230.html, last visited Oct 5, 2024.

ネジメントシステム）[47] と同じ品質マネジメントシステムの系統に属する国際規格である。

なお，プロダクト要件とマネジメント要件をどのように組み合わせるのがよいかについては，試行錯誤が続いている。たとえば，米国におけるソフトウェア医療機器の承認プロセスについて，個々の医療機器（プロダクト）ではなくそれを製造する組織（マネジメント）の要件を重視し，一定の基準を満たす企業には製品の個別承認プロセスを簡素化または省略できるようにする試みが行われた（Pre-certification Program）[48]。しかし結果的には，ソフトウェアの開発プロセスの多様性や，組織評価の基準の欠如，臨床性能やサイバーセキュリティなどデバイス固有の要素を認定する必要性などを理由として，このアプローチは有効に機能しなかったことが報告されている。最近では、プロダクト認証とマネジメント認証を組み合わせたジョイント認証（Joint Certification）と呼ばれる制度の検討も進んでいる[49]。

④　共同規制

AI のように，官民の間での情報の非対称性が拡大している（民間の方が政府よりも圧倒的に多くの情報を有するようになっている）分野においては，政府が一方的に規律内容を決めるアプローチが有効ではない一方で，自主規制に完全に委ねることでは適切にリスクを管理できないという問題がある。そこで，それらの中間，すなわち官民で共同して解決策を導く「共同規制」のアプローチ[50] が有効であると考えられている。

たとえば，内閣府 AI 戦略チームの「『AI 制度に関する考え方』について」では，「技術革新の速さや不確実性に迅速・柔軟に対応するため，ハードローとソフトローの組合せによる『共同規制』の考え方を念頭に，たとえば，リスク対応や情報提供に関する基本的な事項は法令で定め，細則については，AISI（筆者注：AI セーフティ・インスティテュート）等の専門機関において官民共同でガイドラインや規格を策定する方法も考えられる。」としている。

EU の AI 法を含む NLF（上述の製品安全に関する一連の政策パッケージ）も共同規制の考え方を採用している[51]。AI 法では，ハイリスク AI システムが，EU 整合規格に適合する限りにおいて，同法の要件を遵守したものと推定する旨規定するが[52]，これは，民間主導で定める標準に一定の法規範性をもたせたものだ。さらに，ハイリスク以外の AI システムについても，民間が自主的に行動規範を策定することが推奨されている[53]。

なお，ここでの「民間」とは，被規制者だけを意味するわけではない。被規制者の取引相手や消費者，一般市民，学識経験者など，様々な利害関係者（マルチステークホルダー）の立場や知見を統合した規範策定が，共同規制のあるべき姿である。

47) ISO, ISO/IEC 27001:2022, https://www.iso.org/standard/27001, last visited Oct 5, 2024 .

48) U.S. FDA, *The Software Precertification (Pre-Cert) Pilot Program: Tailored Total Product Lifecycle Approaches and Key Findings*, Sep. 2022, https://www.fda.gov/media/161815/download?attachment, last visited Oct 5, 2024.

49) *ISO/IEC NP 25336 Information technology — Artificial intelligence — High-level framework and guidance for the development of conformity assessment schemes for AI systems* (https://standardsdevelopment.bsigroup.com/projects/9024-10625#/section), last visited Oct 16, 2024.
Kimberly Lucy, "*Creating Trust in AI Through Standards: A Management System Approach*" (https://jtc1info.org/wp-content/uploads/2022/06/01_05_Kimberly_ISO_IEC_AI_Workshop-Trust-in-AI-through-MSS_Kim-Lucy.pdf), last visited Oct 16, 2024. なお、この点について拙稿「AI の認証制度をどう設計するか？」（一般社団法人国際経済連携推進センター，2024 年 10 月）も参照。

50) Office of Communications (UK), *Identifying appropriate regulatory solutions: principles for analysing self- and co-regulation*, 2008, https://www.ofcom.org.uk/siteassets/resources/documents/consultations/8466-coregulation/statement/statement-identifying-appropriate-regulatory-solutions---principles-for-analysing-self--and-co-regulation?v=332518, last visited Oct 5, 2024.

51) Francesco Vigna, *Co-regulation Approach for Governing Big Data: Thoughts on Data Protection Law*, Oct. 2022 (https://dl.acm.org/doi/fullHtml/10.1145/3560107.3560117) last visited Oct 5, 2024.

52) EU AI 法 40 条。

53) EU AI 法 69 条。

⑤ 透明性・アカウンタビリティの重視

変化が速く複雑なAIシステムについては，これを設計・開発・運用する事業者と，規制当局を含む外部の主体との間の情報格差が著しく広がることで，事業者が適切に法令を遵守しているのか，AIシステムによってどのようなリスクが生じているのか，それ対して事業者が適切に対応しているか，といったことを外部から判断できなくなる。

そのため，各国では，AIシステムに関して一定の情報開示を義務付けたり，透明化の基準を策定したりしている例が多い。開示の内容は，AIを使っているという事実，出力根拠の説明，開発に用いたデータ，AIの性能と限界など，多岐にわたる[54]。

日本でも，デジタルプラットフォーム取引透明化法において，商品検索結果の表示に係るアルゴリズムの主要な事項を開示することを義務付けている[55]。

さらに，事前に挙動を予測することが困難なAIシステムについて，事故が発生した際に原因を究明するためには，動作に関する記録が残されていることが必要となる。そのため，AIの動作に関する記録を義務付ける例もある[56]。

⑥ 認証・監査

AIシステムは一般的に極めて複雑であるため，前述のような情報開示があったとしても，第三者がそのAIの信頼性を判断することが困難であることが多い。そのような場合には，専門的な第三者が，何らかの信頼を付与する仕組みが必要とされることがある。

典型的なのは，認証（Certification）の仕組みだ。認証とは，独立した機関が，対象となる製品，プロセス，サービス，またはシステムが特定の要件を満たしていることを証明する文書（証明書）を提供することをいう[57]。

EUのAI法では，ハイリスクAIシステムのプロバイダが，EU整合規格への適合性審査を行うこととされており，その中でも欧州整合法令において適合性評価が義務付けられているものについては，第三者による認証が必要となる[58]。日本でも，AIシステムを含むプログラム医療機器について，専門機関が第三者認証を与える仕組みがある[59]。また，前述の「『AI制度に関する考え方』について」等の政府文書では，第三者認証の導入可能性について度々言及されている。

認証は，一定の基準（標準）に基づいて実施されるものであるが，AIについてはそのリスクが多様で技術変化も速いことから，そのような基準が策定されていないことが多い。その場合に，AIモデルの性能やAIサービスの信頼性を確保するための有力な手段となるのが，専門的な第三者による監査の仕組みだ。たとえば，ニューヨーク市の自動雇用決定ツール法[60]では，採用時に使われるアルゴリズムのバイアスについて，年に1回の独立した第三者によるアルゴリズム監査を受けることが義務付けられている[61]。また，結果的には廃案となったが，カリフォルニアAI安全法[62]でも，高度なAIモデルの開発者が，同法の義務遵守について，年に一度の監査を受けることを義務付けていた。

⑦ 継続的な評価と改善

AIシステムは，流通開始後も継続的にアップデートされ，それに伴ってシステムの信頼

54) 詳細については，AI法研究会国際交流部会「AIの透明性に関する国際的なルール比較」及び「透明性比較表」（2024年7月）（https://www.aiandlawsociety.org/paper，2024年10月5日最終閲覧）参照。

55) デジタルプラットフォーム取引透明化法5条2項1号ハ。

56) EU AI法12条。なお，道路運送車両法41条2項は，自動運行装置に作動状態の記録を義務付けるが，これはAIの作動内容の記録ではなく，AIが作動していたかどうかの状態を確認するものにすぎない。

57) ISO, *Conformity Assessment* (https://www.iso.org/conformity-assessment.html), last visited on Oct 16, 2024.

58) EU AI法6条1 (b)，43条。

59) 前掲注40

60) Chapter 5 of Title 6 of the Rules of the City of New York, Subchapter T: Automated Employment Decision Tools, 2023, https://rules.cityofnewyork.us/wp-content/uploads/2023/04/DCWP-NOA-for-Use-of-Automated-Employment-Decisionmaking-Tools-2.pdf, last visited Oct 5, 2024.

61) 同法§ 5-301

62) カリフォルニアAI安全法22603 (e)

性や安全性にも変化が生じるため，流通開始後にもリスクをモニタリングしたり，監査を受けたりすることが義務付けられる場合がある。たとえば，EU の AI 法では，ハイリスク AI システムのプロバイダが，AI 技術の性質及びリスクに応じて，流通後のモニタリングシステムを確立し，文書化しなければならないと定める[63]。⑥で記載した，ニューヨーク市の自動雇用決定ツール法やカリフォルニア AI 安全法（廃案）が求める年に一度の AI 監査も，同様の継続的モニタリングの必要性から求められているものである。

⑧　国際的な協調

AI サービスは容易にグローバル展開されるために，以上の様々な項目についても，国際的な連携が不可欠になる。現時点での国際連携の進捗については別稿[64]で分析したので詳細は割愛するが，G7 を含む主要国は，AI 規制における国際連携の重要性を強調しており，標準の策定や規制の枠組[65]等に関する国際協力が進められている。

Ⅲ．「AI 規制論」のスコープに関する疑問

以上，AI 規制について現在各国で議論されていることを概観した。しかし，これらの議論は，本当に「AI」という特殊な文脈の議論と位置付けるべきなのだろうか。本節では，この点について，① AI と人間のリスクの相対性，② AI と人間のリスクの不可分性，③ AI 規制論の根拠の一般的妥当性，という観点から検討する。

1　AI と人間のリスクの相対性

第一の問いは，「なぜ AI のリスクだけを特別視する必要があるのか？」ということである。本稿Ⅱ 1 において，AI がもたらすリスクを概観した。その項目を見ると，自動車事故やプライバシー侵害のような個人レベルのものから，性・人種差別のようなコミュニティレベルのもの，デマの拡散による選挙の妨害といった国家レベルのもの，そして気候変動のようなグローバルレベルのものまで，AI はあらゆるリスクの元凶であるようにも見える。しかし，これらのリスクは，少なくともカテゴリーとしては以前から存在するものであり，AI の登場によってはじめて問題となったリスクではない。

ここで，改めて AI の定義についてみてみたい。責任ある AI の実装に向けて G7 諸国を含む多くの先進国が遵守を約束している OECD AI 原則では，AI システムを以下のように定義している[66]。

「明示的または暗黙的な目的のために，受け取った入力から予測，コンテンツ，レコメンド，物理的または仮想環境に影響を与え得る決定の出力を生成する方法を推測する機械ベースのシステム。実装後の自律性と適応性のレベルは，個別の AI システムによって異なる。」

この中でも特に重要なのは，「受け取った入力から」「出力を生成する方法を推測する」「機械ベースの」システム，という点であろう。すなわち，入力と出力をつなぐ処理を，データから学習する「機械学習」を行うのが，ここでいう AI である。つまり，人間が予め計算式を決めておくのではなく，機械がデータを学習して最適な計算式を作り出すということである。機械によるアルゴリズムの導出は，学習したデータを統計的に分析することによって行われる。そして，導出（学習）されたアルゴリズムに対してある入力を行うと，その入力に対して確率に基づく出力が得られる。

簡単な「機械学習」プログラムであれば，「機械」ではなく人間が電卓（もっといえば

63) EU AI 法 72 条。
64) 前掲注 13
65) 欧州評議会 AI 条約参照。
66) OECD, The OECD Framework for Classifying AI systems, at 5, https://wp.oecd.ai/app/uploads/2021/06/OECD-Framework-for-Classifying-AI-Standard-deck.pdf, last visited Oct 5, 2024.

紙と鉛筆）を使ってもできる。ただ，それでは現実の複雑な事象を予測できるだけの精度が出ない。しかし，社会のデジタル化によって膨大なデータと莫大な計算資源が獲得可能になったことで，幾層にも重なる関数を用いた「ディープラーニング」の実装が可能になったことが，ブレークスルーとなった。ディープラーニングでは，その演算が極めて複雑になるために，人間が同じ計算をしたりアルゴリズムの意味を理解したりすることは現実的には不可能である。しかし，それは，機械学習の根本的な原理が変わったことを意味するわけではない。あらゆる複雑性を排して敢えて端的に表現すれば，AIとは，「超高性能統計確率計算機」である。

そしてここで重要になるのは，一定のデータから統計と確率に基づく判断を行うこと自体は，これまでの社会で人類が行ってきたものだということだ。販売データに基づく消費者の動向予測や，過去の機械の故障データに基づく部品交換時期の予測といったビジネス上の判断はもちろんのこと，たとえば顔の認識ですら，それまで各人が認識した顔データから一定の傾向を看取する学習を行ったものと理解することもできる（その結果として，たとえば多くの日本人にとっては，アジア系の人物に比べてアフリカ系の人物の識別の方が困難である，といったバイアスが生じる）。

そうだとすれば，AIがこのような統計・確率に基づく判断を高精度かつ高速に実施できることの何が問題なのだろうか。

もちろん，AIがこれらのリスクを増幅させることは否定できない。たとえば，自律飛行ドローンのプログラムに欠陥があれば何千機というドローンの墜落につながるかもしれないし，採用判断に使われるAIに不適切なバイアスがかかっていれば，その差別は数万人の採用判断に及ぶかもしれない。また，生成AIは，巧妙な偽動画やコンピュータウィルスなど，従来は専門的な技術や知識がなければできなかった有害コンテンツの製作コストを圧倒的に下げ，犯罪のリスクを増大させてきた。

しかし，こうしたAIによるリスクの拡大は多くの場合，量的な差であって，質的な差ではない。すなわち，「害悪（リスク）が大きいから規制すべき」なのであって，「AIだから規制すべき」なのではない。そうだとすればAIという特定の技術のみをターゲットとして規制を行うのではなく，害悪の結果に着目して行うことを原則とすべきではないか。

このような指摘は，規制の「技術中立性」の問題として議論されてきたことである。規制の技術中立性とは，規制はその目的を達成するために，特定の種類の技術の使用を強制したり，優遇したりすべきではないという原則であり[67]，国連のほか，日本[68]，EU[69]，米国[70]，英国[71]等の諸国における規制の基本原則となっている。

67) UNESCAP “Explanatory note to the Framework Agreement on Facilitation of Crossborder Paperless Trade in Asia and the Pacific”, para.20, Nov. 2016, https://www.unescap.org/sites/default/files/Revised Explanatory Note_Post 2nd IISG Version_0.pdf, last visited Oct 5, 2024.

68) 令和5年6月規制改革推進会議答申において，「特定の技術・手段などを求める画一的で『事前型の規制・制度』から，技術中立的でリスクベース・ゴールベースの柔軟な『事後型の規制・制度』への見直しを進めていかなければならない」と述べられている。規制改革推進会議「規制改革推進に対する答申～転換期におけるイノベーション・成長の起点～」2頁（2023年6月）（https://www8.cao.go.jp/kisei-kaikaku/kisei/publication/opinion/230601.pdf，2024年10月5日最終閲覧）。

69) 欧州委員会ウェブサイト“*Technology Neutrality*”（https://joinup.ec.europa.eu/collection/common-assessment-method-standards-and-specifications-camss/solution/elap/technology-neutrality#:~:text=It%20guarantees%20freedom%20of%20choice,nor%20discriminate%20against%20any%20technology，2024年10月5日最終閲覧）。なお，規制の技術中立性を原則としているEUにおいて，なぜAIという技術非中立な規制が正当化されるのかについて，明確な説明はない（参照：European commission, Artificial Intelligence – Questions and Answers, Why do we need to regulate the use of Artificial Intelligence?, https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683, last visited Oct 5, 2024）。

70) President William J. Clinton & Vice President Albert Gore, Jr., *A Framework For Global Electronic Commerce*, II. 3., https://www.w3.org/TR/NOTE-framework-970706, last visited Oct 5, 2024.

71) Azad Ali et al. , *UK Regulators Publish Approaches to AI Regulation in Financial Services*, May 2, 2024, https://

このように，AI という技術のみに着目した規制を議論することは，AI と人間の機能の相対性という観点からも，また規制の技術中立性という一般原則からも，疑問がある。

2　AI と人間のリスクの不可分性

実社会において，人間と AI は単独に作用するのではなく，常に相互に補完しながら機能している。分かりやすい例は，生成 AI である。そこでは，人間のインプットに対して AI が何らかの文書や画像，動画などを生成して回答し，それに対して更に人間が指示を与えることで，より洗練された回答が導かれる。このような相互のやりとりを通じて生成されたコンテンツが，社会に何らかの影響を与える。

医療における AI によるレントゲン画像診断や，採用における AI によるエントリーシートのレビューなども，AI の回答がそのまま確定診断や採用判断に採用されるわけではなく，人間が最終的な決定を行うことがほとんどである。

翻って考えれば，現代社会において，AI の判断の影響を受けない人間の行動はほとんどない。情報の検索，移動経路の導出，購買する商品の選定，そしてソーシャルネットワークなど，我々が日々用いるあらゆるシステムには AI が組み込まれている。また，電力供給システムや道路交通システムなど，日々の生活を支えるのに欠かせないインフラにも AI が活用されている。他方で，上述のように，AI の判断のみによって我々の判断やリスクが完全に決定される場面も少ない。AI は，人間が上書きした判断の集合から学習し，アルゴリズムをアップデートする。

つまり，AI は人間の認識に影響を与え，人間はその認識に基づき行為し，AI はそのデータから再度学習してアルゴリズムを更新するという循環の中で，人間と AI は相互に浸潤しながら社会活動を行い，同時にリスクを生んでいるのである。そして，このような循環の中で，厳密に「AI のリスク」と「人間のリスク」を線引きできるかどうかは疑わしい。

たとえば，自動運転を原則としつつ，緊急時には人間に運転を引き継ぐ「レベル 3」の自動運転について，次のような事例を考えてみよう。自動運転のアルゴリズムに何らかの不具合が生じたため，システムが人間に運転を引き継ぐためのアラートを発出したが，運転者の注意が散漫になっており，アラートへの対応が遅れ衝突事故を起こしてしまった。一見すると，自動運転アルゴリズムの欠陥及び注意が散漫になっていた運転者の落ち度であるといえそうである。しかし，全てのリスクシナリオを想定した完全な自動運転アルゴリズムを構築することは不可能であるし，問題があった場合にアラートを出したこと自体はアルゴリズムの正常な作動ともいえるため，一体何をもって「アルゴリズムの欠陥」といえるのかが明らかではない。また，自動運転アルゴリズムに問題があった場合に，単に人間に運転を引き継ぐのではなく，予備システムによって安全な路肩に駐車する機能も搭載可能であったかもしれない。さらには，自動運転モードにおいて人間の注意が散漫になりがちであることを踏まえ，運転者の集中度をシステムが監視して注意散漫時に警告を行うこともできただろう。これらのシステムが備わっていれば，運転者は事故を防げたかもしれない。ただでさえ，自動運転モードでは人間の注意は散漫になる。このような状況で，自動運転システムの欠陥または人間の過失のいずれかが事故の原因だと特定することは正当だろうか。

このように，実態としての事故は AI と人間が一体として作用する結果として生じるのであり，自動運転アルゴリズム（すなわち AI システム）の問題か運転者（人間）の問題か，と二元的に整理できる問題ではない。

もう一つ，生成 AI によるリスクの典型例とされる，政治に関するディープフェイク（たとえば，首相があたかも卑猥な発言をし

www.skadden.com/insights/publications/2024/05/uk-regulators-publish-approaches-to-ai#:~:text=The%20PRA%20and%20FCA%20both,mandate%20or%20prohibit%20specific%20technologies, last visited Oct 5, 2024.

ているかのようなフェイク動画が広まった事例[72]）について考えてみよう。たしかに，そのような不適切な動画を容易に作り出せてしまう生成AIにも問題はある。しかし，そもそもの元凶は，悪意をもってそのような動画を作り出した人間にある。また，そのような動画を拡散するソーシャルメディアこそがリスクを爆発的に増加させているともいえる。ソーシャルメディア上でそのようなセンセーショナルなコンテンツが拡散される背景には，人間（ユーザー）がもつ本能的な承認欲求や，クリック数に応じて報奨金がもらえるソーシャルメディアのエコシステムなども関係している。

以上のように，AIが関連するリスクは，AI自体のリスクというよりも，AIと人間，そしてその他のシステム（他のアルゴリズムやハードウェア，経済システムや環境等）が相互に影響し合うことによるシステミックなリスクとして理解する必要がある。そのような現実を踏まえた場合に，AIという特定の技術要素のみを取り出して規制の対象とすることは，リスクの本質を見逃し，社会にとって利益よりも多くのコストをもたらすことになりかねない。

もちろん，大きなシステムの一要素としてのAIの機能（自動運転車のアルゴリズムを構成する画像認識システムの精度など）について，何らかの要件を課す必要がある場合もあるであろう。しかし，これはあくまで規制の一要素を構成するモジュールに過ぎず，制度全体の設計にあたっては，個別のAIシステムを対象とするのではなく，人間や環境との相互作用を踏まえて，AI，その他のアルゴリズムやハードウェア，組織内のルールや体制，そしてユーザの属性等を考慮したシステム全体としてのリスクを評価する必要があると考えられる。

3　AI規制のアプローチの非独自性

以上，AIという技術がこれまでの人間の判断の延長にあること，また現実世界において人間とAIのリスクを明確に分離できないことを示した。以下ではその帰結として，本稿Ⅱで述べたようなリスクやそれに対応するためのアプローチが，AI以外の文脈にも共通するものであることを示す。

まず，Ⅱ1で挙げた「AIリスク」の項目は，個人レベル（生命・健康，プライバシー，有形・無形の財産権，自律性，雇用，その他の基本的人権等），コミュニティレベル（差別，多様性の喪失等），国家レベル（安全保障，民主主義，法の支配等），そしてグローバルレベル（気候変動，力の集中と地域格差等）に至るまで，AIに固有のものではない。また，「AI原則」としてまとめられる公平性，プライバシー，安全性，セキュリティ，透明性，アカウンタビリティといった価値も，AIを使わずとも尊重されるべき重要な価値である。

それでは，AI規制の手法についてはどうであろうか。Ⅱ4では，これを①リスクベースアプローチ，②ハードローによる大枠の設定とソフトローによる具体化，③プロダクト要件とマネジメント要件の組み合わせ，④共同規制，⑤透明性・アカウンタビリティの重視，⑥認証・監査，⑦継続的な評価と改善，⑧国際的な協調，という要素に分解したが，これらの要素も，他の規制分野でこれまで議論されてきたことと大きく重なっている。

まず，①リスクベースアプローチは，監督・立法・内部統制といった複数の階層において，リスクの高い分野に重点的にリソースを振り向けるべきという原則であったが，これはAIに限った話ではない。AIの文脈では，リスクの範囲が多岐にわたることから，とりわけこの用語が用いられる傾向があるが，既にみたように，規制監督におけるリスクベースの考え方自体は金融規制の分野において定着していたし，立法におけるリスクベースは比例性の原則として理解できる。さらに，内部統制の文脈でも，とりわけ監査の文脈において，「リスクベースアプローチ」は20世紀

72) 日本経済新聞「岸田文雄首相の偽動画拡散　生成AIか，日テレロゴも」（2023年11月4日）（https://www.nikkei.com/article/DGXZQOUE042850U3A101C2000000/，2024年10月5日最終閲覧）。

のうちから導入されている[73]。

②ハードローによる大枠の設定とソフトローによる詳細の補完についても，AI 以外の文脈においてその重要性が指摘されてきた。

ハードローだけでは時代の変化に対応できないことは，とりわけ変化の速い金融やデジタルの分野などで指摘されている。その際，規制は達成手段ではなく結果を基準にして行うべきとする「結果（Outcome）ベースの規制」の考え方が，英国や日本を含む諸国において重視されている[74]。

他方，ソフトローは，1980 年頃，国際法の文脈における人権や環境の分野で注目を集め[75][76]，コーポレートガバナンス[77]や知的財産[78]，そして規制一般[79]など，広範な分野でその重要性を認識されている。ソフトローの特長は，リスクや環境の変化に迅速に対応しやすい点にあるが，その柔軟性が求められるのは，AI 分野だけではない。そして，そのソフトローを民間の知見を取り込みながら策定するのが，以下④で述べる「共同規制」のアプローチである。

③プロダクト要件とマネジメント要件の組み合わせについても，AI の登場以前から活用されていた規制枠組である。たとえば，プログラム医療機器について，各国で医療機器自体の要件と製造者の品質マネジメントシステムの要件の双方が求められてきたし，消費生活用製品などについても，製品自体の安全性のほかに，品質管理システムの整備が求められてきた[80]。マネジメント要件としては，EU の一般データ保護規則（GDPR）のように，一定の場合に責任者（データ保護監督者）を置くことを求める例もある[81]。

④官民で共同して解決策を導く共同規制についても，近年様々な分野で注目を集めている。EU では，今世紀初頭から共同規制の活用が提唱されており[82]，現在では，AI のみならず，デジタルプラットフォーム規制やデータ規制など様々な分野で共同規制のアプローチが採用されている[83]。日本でも，デジタルプラットフォーム規制を含むデジタル規制全般において，共同規制の重要性が指摘されている[84]。

⑤事前規制よりも透明性やアカウンタビリ

73) Financial Reporting Council, Internal Control: Revised Guidance for Directors on the Combined Code, 1999, revised in 2005, https://www.accountancyeurope.eu/wp-content/uploads/FRC_Revised_Guidance_for_Directors_on_Combined_Code_051027102005121030.pdf, last visited Oct 5, 2024.

74) UK Financial Conduct Authority, Our Strategy 2022 to 2025, at 5, 2022, https://www.fca.org.uk/publication/corporate/our-strategy-2022-25.pdf, last visited Oct 5, 2024. 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「GOVERNANCE INNOVATION――Society5.0 の実現に向けた法とアーキテクチャのリ・デザイン――」36 頁（2020 年 7 月 13 日 ）（https://www.meti.go.jp/shingikai/mono_info_service/governance_model_kento/pdf/20200713_1.pdf，2024 年 10 月 5 日最終閲覧）。

75) Google Books Ngram Viewer, ‘soft law’, https://books.google.com/ngrams/graph?content=soft+law&year_start=1800&year_end=2022&corpus=en&smoothing=3, last visited Oct 5, 2024.

76) 齋藤民徒「ソフトロー論の系譜：国際法学の立場から」（2005 年 7 月）（https://www.j.u-tokyo.ac.jp/coelaw/COESOFTLAW-2005-7.pdf，2024 年 10 月 5 日最終閲覧）。

77) Financial Reporting Council, UK Corporate Governance Code, Jan. 2024, https://www.frc.org.uk/library/standards-codes-policy/corporate-governance/uk-corporate-governance-code/, last visited Oct 5, 2024.

78) 内閣府知的財産戦略推進事務局「ソフトローの活用について」（2021 年 4 月 16 日）（https://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kousou/2021/dai6/siryou3.pdf，2024 年 10 月 5 日最終閲覧）。

79) European Commission, European Governance A White Paper, Jun. 25, 2001, https://ec.europa.eu/commission/presscorner/detail/en/DOC_01_10, last visited Oct 5, 2024.

80) 経済産業省「製品安全に関する事業者ハンドブック」（2012 年 6 月）（ https://www.meti.go.jp/product_safety/producer/jigyouhandbook.pdf，2024 年 10 月 5 日最終閲覧）。

81) EU 一般データ保護規則（GDPR）37 条。

82) European Commission, *supra* note 70.

83) 生貝直人「ソフトローを活用したルール形成と自主・共同規制」（2021 年 4 月 16 日）（https://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kousou/2021/dai6/siryou4.pdf，2024 年 10 月 5 日最終閲覧）。

84) デジタル庁「デジタル原則に照らした 規制の一括見直しプラン」（官民連携原則）2-3 頁（2022 年 6 月 3 日）（https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/cb5865d2-8031-4595-8930-8761fb6bbe10/e3650360/20220603_meeting_administrative_research_outline_07.pdf，2024 年 10 月 5 日最終閲覧）。

ティを重視するアプローチは，官民間の情報の非対称性の増大に伴い定着してきた。リスク状況やステークホルダーの関係が複雑な金融，人権，環境といった分野では，いずれも開示が重要な規制手段とされてきたし[85]，AI以外のデジタル分野においても，透明性は主要な規制手段とされている[86]。

⑥認証・監査についても，上記③のプロダクト要件とマネジメント要件の双方について，ハイリスク分野を対象に，AIの台頭以前から認証や監査が義務付けられる場合があった。認証については，既に紹介した自動車や医療機器，消費生活用製品等に関するものはいずれも，AI以外のシステムにも義務づけられるものである。また，法的義務ではないが，情報セキュリティマネジメントシステムの認証については，国内外で多くの企業が認証を取得している。監査については，たとえば，米国の医療保険の携行性と責任に関する法律（HIPPA）において，電子的な個人情報を含むハードウェア，ソフトウェア，手続き，アクセスの記録等の活動の監査が求められている[87]。

⑦継続的な評価と改善については，これまでも，「市販後監視」（Post-market surveillance）として医療や食品分野を中心に義務付けが行われてきた。たとえば，医療機器に関する国際標準であるISO13485:2016には市販後監視が規定されており，これに整合する日本の薬機法[88]やEUの医療機器規則[89]，米国の連邦食品・医薬品・化粧品法[90]において，それぞれ販売後の製品の評価と改善が求められている。インシデントの共有については，消費生活用製品安全法における重大製品事故の報告義務[91]や，米国の消費者製品安全法（Consumer Product Safety Act）における報告義務[92]などの例がある。

⑧国際的な協調については，経済・社会のグローバル化が著しく進んだ現在において，その重要性は論を俟たない。デジタル，環境，人権等，幅広い分野で国際的な連携の取組が進んでいる。

Ⅳ. なぜ「AI規制論」を「一般規制論」へ昇華すべきなのか

前節では，AIの技術的性質やリスクの特徴を踏まえ，AIのみを規制の対象として括りだすことへの疑問を示すと共に，実際に「AI規制のアプローチ」として語られるリスクや規制手法が，AIのみならず様々な分野で既に適用されているものであることを示した。

しかし，それだけでは，「AIのために特別な規制を行う必要がない」ということを示したに過ぎず，「あらゆる規制にAI規制のアプローチを取り込むべきである」という本稿の主張を十分に示したことにならない。そこ

85) 人権について：国際連合人権理事会「ビジネスと人権に関する指導原則：国際連合『保護，尊重及び救済』枠組実施のために」（UNGP）21項（2011年3月21日）（https://www.unic.or.jp/texts_audiovisual/resolutions_reports/hr_council/ga_regular_session/3404/，2024年10月5日最終閲覧）。
環境について：環境省「気候関連財務情報開示タスクフォース（TCFD）」（https://www.env.go.jp/policy/tcfd.html，2024年10月5日最終閲覧）。

86) たとえば，2022年に施行されたデジタルプラットフォーム取引透明化法は，その名のとおり，大規模なデジタルプラットフォームに対して事前規制を課すのではなく，取引条件の開示や公正な取引を行うための措置等の透明化を求める内容である。

87) Health Insurance Portability and Accountability Act (HIPPA), 45 CFR § 164.312(b).

88) Federal Food, Drug and Cosmetic Act (FDCA), Section 522.

89) Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, Arts. 83-86.

90) Consumer Product Safety Act (CPSA), Section 15 (b).

91) 消費生活用製品安全法35条。

92) United States Consumer Product Safety Commission, "*Duty to Report to CPSC: Rights and Responsibilities of Businesses*" (https://joinup.ec.europa.eu/collection/common-assessment-method-standards-and-specifications-camss/solution/elap/technology-neutrality#:~:text=It%20guarantees%20freedom%20of%20choice,nor%20discriminate%20against%20any%20technology), last visited Oct 16, 2024.

で，本節では，なぜこれまで「AI 規制論」の文脈で論じられてきたアプローチを，一般的な規制論に展開させるべきなのかについて論じる。

1　AI の実装領域の拡大

本稿Ⅲ 1 で述べたように，AI の機能は，従来人間が行っていた統計的・確率的な作業を高速で反復・増幅することであり，今後あらゆる分野での活用が見込まれる。また，Ⅲ 2 で述べたように，AI と人間のリスクを明確に分離することは不可能である。すなわち，我々の社会は，いかなる分野においても AI リスクと切り離せない社会になっていく。したがって，AI 規制のアプローチが，AI システム一般に妥当するものであるとすれば，今後あらゆる分野において同様のアプローチを導入する必要が出てくるだろう。

2　社会システム全般における不確実性の増大

「AI 規制のアプローチ」の根拠となる AI システムの特徴は，実は AI 以外のシステムにも当てはまる。

Ⅱ 4 では，AI 規制のアプローチの根拠として，(i) 予測不可能性・説明不可能性，(ii) バリューチェーン上の主体の多さ，(iii) 技術革新や普及の速さ，(iv) 外部からの信頼性判断の難しさ，(v) グローバル化，といった課題を挙げた。たしかにこれらは，AI システムが一般的に備える特徴とはいえるものの，AI システムに固有の特徴ではない。

たとえば，(i) 予測不可能性・説明不可能性については，AI（ディープラーニング）システム自体の複雑性に基づく予測不可能性や説明不可能性が課題であるといった説明がなされることも多い。しかし，人間の行動や自然現象はそもそも予測不可能かつ説明不可能なことが多く，さらにそれらがグローバル化やデジタル化に伴って反復・増幅されている。ビジネスの世界では，現代社会を，不確実性 Volatility（変動性）・Uncertainty（不確実性）・Complexity（複雑性）・Ambiguity（曖昧性）の頭文字をとって，「VUCA」社会と表現することもある[93]。AI は確かに不確実性の一因となってはいるが，AI システムが関与せずとも世界は予測不可能・説明不可能な事象に溢れている。

(ii) バリューチェーン上の主体の多さとそれに伴う課題についても，AI に限った話ではない。たとえば，ファッション業界では，素材生産，流通，デザイン，縫製，販売などの過程で多くの主体が関与しており，それぞれの役割を担う主体が複数国にまたがっているため，バリューチェーン全体での環境や人権への影響のコントロールが喫緊の課題となっている[94]。また，AI に限らないデジタルサービス一般に目を向けても，そこにはハードウェアの設計者・製造者，クラウドサービス提供者，通信事業者，ＯＳ（オペレーティングシステム）の開発・提供者，デジタルプラットフォームサービスの提供者，データの提供者，ユーザー（事業者または消費者）など多様なステークホルダーが混在しており，いずれにせよ各サービス間の透明性や責任の分配が問題となる。

(iii) 技術革新や普及の速度についても，AI に限られる話ではない。過去 20 年ほどを振り返っても，スマートフォンやソーシャルメディア，クラウドコンピューティング，再生可能エネルギーなど，様々な分野におけるイノベーションとその実装が爆発的に進められてきた。

(iv) 外部からの信頼性判断の難しさについても，社会の複雑化に伴う情報の非対称性の拡大によって，AI のみならず，システムから組織まで様々なレベルで問題となっている。既にみたように，人権・環境・コーポ

93) Nate Bennett & G. James Lemoine, *What VUCA Really Means for You*, Jan.-Feb. 2014, HARVARD BUSINESS REVIEW.

94) Gian Bonanni et al. , *Explainer: What is fast fashion and how can we combat its human rights and environmental impacts?*, Australian Human Rights Institute, https://www.humanrights.unsw.edu.au/research/commentary/explainer-what-fast-fashion-human-rights-environmental-impacts, last visited Oct 5, 2024.

レートガバナンスなど多くの分野で透明性の規律が重視されるようになってきているのも，こうした課題の顕れであるといえる。

最後に，(v) グローバル化が AI 以外の領域で進んでいることは，言うまでもない。

このように，AI 規制のアプローチの根拠とされるリスク状況の変化は，実は AI に固有のものではなく，デジタル化やグローバル化に伴う社会システムの複雑化，産業構造の変化，そして不確実性の増大という，より大きな文脈の中で生じている変化である。したがって，AI 規制のアプローチは，AI という狭い文脈の中でのみ議論されるべきではなく，社会全般に関するアジェンダとして議論されるべきである。

3 「AI 規制のアプローチ」による法の支配の実質化

AI 規制論を一般的な規制論として展開すべきという本稿の主張を裏付けるより本質的な根拠は，「AI 規制のアプローチ」こそが，現代社会において「法の支配」を実質化する手段であることである。

法の支配については様々な定義があるが，国連が 2004 年に発行した報告書は，これを以下のように定義している [95]。

「(法の支配とは,) 公的・私的を問わず，すべての個人，機関，団体（国家を含む）が，公布され，平等に執行され，独立に裁定され，国際的な人権規範と基準に合致する法律に対して責任を負う統治原則を指す。さらに，法の支配は，法の優位性，法の下の平等，法へのアカウンタビリティ，法の適用における公平性，権力分立，意思決定への参加，法的安定性，恣意性の回避，手続的・法的透明性といった原則を遵守するための措置も意味する。」

近代の統治モデルも，このような法の支配を実現するために整備されたものであるが，デジタル化やグローバル化によって著しく複雑化し加速した現代社会では，様々な困難に直面している。

ローレンス・レッシグが，人々の選択肢や行動範囲がデジタルシステムの構造（アーキテクチャ）によって規律されるようになる，と説いたのは 1999 年だが [96]，それから四半世紀を経て，デジタルシステムがもたらす影響力（Power）は，当時と比較にならない程大きくなった（Ⅲ 2 参照）。この新たな力は，変化が速く複雑で，グローバル化されている。それは，既存の法の規律が十分に及ばない領域の拡大を意味する。

システムに関する情報の非対称性の拡大によって，これを設計・運用していない者には，それがどのように機能しているのか理解できず，それによって誰にどのような影響が生じているのかも理解できない。その上，相手方が国外の主体であればそもそも立法や執行力の管轄が及ばない可能性もある。このことは，「法の優位性」ではなく「人の優位性」――すなわち，特定の人（実際にはその集合としての組織，なかんずく巨大企業ないし統制国家）が恣意的に影響力を行使できる余地――の拡大をもたらす。

こうした事態に対処するためにはどうすればよいのだろうか。

矢継ぎ早に規制を作っていけばよいかというと，必ずしもそうではない。日本の場合，法律の策定には短くとも 2-3 年程度かかることを踏まえると，法律が技術の変化に追いつけない問題を解消することはできない。そのため，法の支配を実質化するためには，法律を，最終的に達成されるべき価値や原則を記述する，結果ベースかつプリンシプルベースのものとせざるを得ない。

そのような状況下では，法の支配が要求する「法へのアカウンタビリティ」も，単に法律の定める行為義務を遵守しているというだけでは不十分であることになる。すなわち，リスクをもたらすシステムの設計・運用主体（民間主体か公的主体かは問わない）が自ら，どのようにして法の定める結果や原則を達成すべきかを決断して宣言すること，そしてそ

95) United Nations Secretary-General, *The Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies : report of the Secretary-General*, ¶ 6, S/2004/616, U.N. DOC., Aug. 23, 2004.

96) LAWRENCE LESSIG, CODE: AND OTHER LAWS OF CYBERSPACE (2000).

の結果について責任を負うことこそが，法に対するアカウンタビリティの尽くし方となる。

各主体によるアカウンタビリティを確保するための要素は，「規範形成」，「モニタリングと評価」，「救済と紛争解決」という 3 つの段階で考えると分かりやすい。

第一の段階は，規範形成である。各主体は，結果ベース・プリンシプルベースの規制の下で，様々な価値を現実のシステムに実装し，ステークホルダーの理解を得るべきである。しかし，これは容易ではない。とりわけスタートアップや中小事業者にとってはそのような専門知識や人材を獲得することが困難である。そのため，マルチステークホルダーの協力の下，柔軟な標準やガイダンスなどが公共財として提供されることが望ましい（共同規制）。このようなプロセスを経ることによってこそ，法の支配が要求する「意思決定への参加」が実現されるのであり，また規範が一定程度明確化されることで，「法の適用における公平性」や「法的安定性」が達成されると共に，「恣意性の回避」も確保される（この点が，政府による一方的な「行政指導」による統治との違いである）。

第二の段階は，モニタリングと評価である。各主体が，自らの宣言を遵守しているか，またそれによって法目的が達成されているかについては，外部から評価することが非常に困難だ。そのため，透明性の確保をベースとしつつ，リスクの大きさに応じて，独立かつ専門的な第三者が認証したり監査したりすることが重要となる。システムのダイナミズムを考慮して，継続的にモニタリングとフィードバックを行っていくことも必要だ。このようなプロセスを経ることでこそ，法の支配が要求する「手続的・法的透明性」が確保される。

第三の段階は，救済と紛争解決である。不確実性が高く事前に明確な行為規範を示すことが困難な現代社会においては，事後的に迅速な被害者への補償や適切な責任分配がなされることが不可欠だ。しかし，多くの要因が相互に作用しており，またリスク間の複雑なトレードオフも存在するシステミックなリスクについては，純粋に法的な知見だけで補償や責任の評価を行うことが難しい。そのためには，法律の専門家だけではなく，経済学，社会学，コンピュータサイエンス，心理学，環境学等を含む様々な専門性を有する中立な第三者が判断材料を提供できる仕組みを整備する必要があるだろう。こうした手続によってこそ，法が「平等に執行され，独立に裁定され」る状況を担保できる。

このように，複雑で不確実性の高いダイナミックな社会においては，国家による立法・行政・司法だけでなく，各システムに関する規範形成，モニタリング・評価，救済・紛争解決という局面（これらは概ね，国家レベルでいう立法・行政・司法の三権分立に対応する）において，マルチステークホルダーによるアジャイルなプロセスを導入し，チェックアンドバランスを適切に機能させることこそが，法の支配の核心としての「権力の分立」の姿である。そして，本稿で整理した「AI 規制のアプローチ」は，そのような新たな法の支配の姿を志向するものだ。

4 情報技術を活用した「AI 規制のアプローチ」の実現可能性

このような「AI 規制のアプローチ」を一般的に展開するにあたっては，統治に必要な情報量が爆発的に増加することが課題となる。抽象的な法規制を実務に落とし込むための具体的なガイダンスや標準の策定にあたっては，サービス提供側，ユーザー側，消費者側など，多様な立場から寄せられる大量の見識や意見を総合的に考慮する必要がある。また，現実社会で生じている生命・健康やプライバシー，そして民主主義などに対する数値化困難なリスクや便益を適切に測定しなければならない。策定後も，それによって目的どおりの効果が生じているか，うまくいっていない点があるとすればなぜかを継続的に評価し，必要に応じて規範を更新する必要がある。このことは，参照すべき情報の爆発的な増加及びダイナミック化を意味し，規制者と被規制者の双方に多くのコストをもたらす。

これまでの社会において，こうした多量の

情報処理は，政策担当者にとっても事業者にとっても事実上極めて困難であった。その結果，たとえば日本では，有識者を集めた審議会形式で規制やガイドラインの方針を定め，事業者はそれに形式的に従う，ということが一般的に行われてきた。

しかし，情報技術の発展は，これまで困難であったあらゆるデータに基づく統治を可能にする。たとえば，リスク評価において，プライバシーに関するデータにユーザーがどれだけの受容性を有しているかを計測したり[97]，ソーシャルメディアにおけるアルゴリズムがユーザーの政治的意見にどのような影響を与えているかといった事象を計測したり[98]することも，以前より遥かに容易になった。

多様な意見の集約と論点の抽出も，生成AIなどを使えばより効率的に実施可能となるだろう。台湾の合意形成プラットフォームである”vTaiwan”では，ライドシェアの許可などの特定の政策アジェンダに対する人々の意見をアルゴリズムでグループ分けして，世論の可視化や論点の抽出に活用しており，同様の取組は，フランスやニュージーランドでも採用されている[99]。

モニタリングにあたって参照すべきベストプラクティスやインシデントについても，データベースに平文で登録しておけば，AIが状況に応じた最適な事例を引用してくれるようなサービスを構築することは可能だろう。リスク状況のリアルタイムな監視は，センサから収集したデータを自動的に分析し続けることで実現できるし，そのような仕組みは既に多くの企業で実装されている。

紛争解決の局面においても，アルゴリズムで過去の判断の傾向を分析したり[100]，AIが調停を補助したり[101]するなど，情報処理や解決の一部をシステムに委ねる試みは始まっている。

このように，法の支配を深化させる「AI規制のアプローチ」は，発達した情報技術を活用してこそ可能になる。もちろん，その際には，その情報技術が信頼に足るものであること，すなわち法の支配に服していることが必要である。このように，法の支配と先端的な情報技術は，相互にその存在を必要とする入れ子構造となっているのである。

Ⅴ．「AI規制論」を「一般規制論」に昇華させることの意義と課題

1 AI規制論を一般規制論へ昇華することの意義

以上のように，AI規制論を一般規制論へ昇華することには，どのような意義があるのだろうか。

第一の意義は，これまでAI規制論として蓄積されてきた知見やノウハウを，他の分野のガバナンスにも応用できることである。Ⅱ2で述べたとおり，基本的人権や民主主義，持続可能性といった基本的価値や，安全性やプライバシー，公平性，透明性といった原則を，現代社会においてどのように定義し直し保護していくかという問いについては，「AI原則」や「AI倫理」の問題として世界各国で議論が重ねられてきた。また，多様なス

97) Jaspreet Bhatia et al. , *Empirical Measurement of Perceived Privacy Risk*, Vol. 25, Issue 6, Article 34, ACM TRANSACTIONS ON COMPUTER-HUMAN INTERACTION,(2018).

98) Andrew M. Guess et al., *How do social media feed algorithms affect attitudes and behavior in an election campaign?*, 381 Science,398, 404 (2023).

99) vTaiwan, Where do we go as a society?, https://info.vtaiwan.tw/, last visited Oct 5, 2024.

100) Olga V. Mack, From ‘It Depends’ To Data-Backed Answers: Automating Legal Case Analysis —AI offers a shortcut for optimizing the legal case analysis process—, Sep. 5, 2023, https://abovethelaw.com/2023/09/from-it-depends-to-data-backed-answers-automating-legal-case-analysis/#:~:text=AI%20offers%20a%20shortcut%20for%20optimizing%20the%20legal,answer%20to%20clients%E2%80%99%20what-if%20questions%20beyond%20%E2%80%9CIt%20depends.%E2%80%9D, last visited Oct 5, 2024.

101) 加藤猛「非暴力コミュニケーションのメディエーションに向けた生成AIの試行」（2023年）（https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/284630/1/tec.rep_202308_Kato.pdf，2024年10月5日最終閲覧）。

テークホルダーによる共同規制や複雑なシステムに対する認証などは，AI 規制の策定過程で数々の教訓を得られている。たとえば，米国 NIST による AI リスクマネジメントフレームワーク策定にあたっての大規模なマルチステークホルダープロセス[102]や，シンガポール政府による，AI の信頼性をテストするためのオープンソースシステム（AI Verify）[103]の公表などのアプローチは，AI に限らずあらゆる分野でのガバナンスの実施にあたって参考となるだろう。

第二の意義は，規制の複雑化を避けられることである。AI やブロックチェーン，量子コンピュータなど，新たな技術が登場する度に，屋上屋を架すように規制を重ねていくと，政策担当者も被規制者側も終わりのない執行コストやコンプライアンスコストに苛まれることになる。法律には，社会にとって確保されるべき価値（結果）について技術中立的に規定し，その達成手段については様々なシステムや組織を組み合わせて柔軟かつシステミックに検討していくことが，時代の変化に強い（Future Proof な）規制体系である。

第三の意義は，AI 規制のアプローチによるイノベーションの促進と安全性の確保の両立である。航空機，製薬，労働安全などの分野において，失敗や事故について非難するのではなく，オープンなカルチャーで問題の原因を究明し，改善策を検討するアプローチの方が，結果的に高い安全性を実現できることが実証的に知られている[104]。AI 規制のアプローチに見られるような協調型のガバナンスは，単にイノベーションを阻害しにくいだけでなく，安全性を向上させるものでもあるのだ。

2 AI 規制論を一般化するにあたっての課題

一方で，AI 規制論を一般化するにあたっての課題もある。

第一に，AI 規制論自体が，国内外で現在進行中の議論であり，いまだベストプラクティスと呼べるものが確立されていないことだ。たとえば，日本の AI 事業者ガイドラインは，過去に発行された 3 つのガイドラインを統合して整理した内容であり，リビングドキュメントとして今後継続的に更新されることが宣言されている。AI に関する制度設計を担う内閣府の AI 制度検討会に関しては，議事が非公開とされており，マルチステークホルダープロセスが確保されているとは言い難い。EU の AI 法の下では，義務の実践に関する多くの内容が EU 整合規格に委ねられているが，その整合規格の内容はいまだ明らかでない。ニューヨーク市の自動雇用決定ツール法では，年に一度のアルゴリズム監査が義務付けられているが，法律の曖昧な定義や，監査人やツール開発者，使用企業の役割に関する不明確な責任分担などによって，実効的な監査には至っていないとの指摘もある[105]。このように AI 規制については一定の枠組が整理されてきているものの，その実践にはまだ試行錯誤が必要である。

第二に，Ⅳ 4 で示したように，AI 規制のアプローチを導入するためには，制度だけではなく技術やインフラ，専門人材などをあわせて整備していく必要があるということだ。多種多様なリスクを数値化するための方法論，マルチステークホルダーによる議論を促

102) NIST, *supra* note 23.

103) AI verify foundation ホームページ（https://aiverifyfoundation.sg/，2024 年 10 月 5 日最終閲覧）。

104) Christopher Hodges, *ETHICS IN BUSINESS PRACTICE AND REGULATION*, 2017, https://assets.publishing.service.gov.uk/media/5a7f3f18e5274a2e87db4afc/Prof_Christopher_Hodges_-_Ethics_for_regulators.pdf, last visited Oct 5, 2024, and *Outcome-Based Cooperative Regulation*, 2023, https://www.theregreview.org/2023/01/02/hodges-outcome-based-cooperative-regulation/, last visited Oct 5, 2024. U.S. Federal Aviation Administration, *Safety Culture Assessment and Continuous Improvement in Aviation: A Literature Review*, 2023, https://www.faa.gov/about/initiatives/maintenance_hf/Safety_Culture_Assessment_Continuous_Improvement_in_Aviation.pdf, last visited Oct 5, 2024.

105) Lara Groves et al. , *Auditing Work: Exploring the New York City algorithmic bias audit regime, in* PROCEEDINGS OF THE 2024 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FACCT '24), 1107, 1120 (2024), *available at* https://doi.org/10.1145/3630106.3658959, last visited Oct 5, 2024.

進するための公共プラットフォーム，コンプライアンスを一定の範囲で自動化するツール，リアルタイムのモニタリングを可能にするシステムなど，規制の策定や履行自体に革新的なアプローチを取り入れるには，法学と経済学，社会学，コンピュータサイエンス，心理学，環境学といった様々な分野の連携が必要になる。

第三に，AI 規制のアプローチを実装するためには，社会の一人ひとりの規制に関するマインドセットを転換する必要があるということである。トップダウンではなく水平的な規範作りを，政治的調整ではなくデータに基づくリスク評価を，失敗の非難よりもベストプラクティスの評価を，情報の秘匿より開示を，先例踏襲ではなく未来志向を，専門性の縦割りではなく学際性を，国内に閉じるのではなく国際性を，といったように，新たな規制モデルの構築にあたっては従来のマインドセットを変革していく必要がある。それを成功させることができるかどうかは，各自が，テクノロジーと共生する社会の在り方をいかに「自分事」として考え，参加し，主張できるかにかかっている。

Ⅵ．まとめ

本稿で主張したことの骨子は，以下のとおりである[106)]。

過去数年間にわたって，国内外における AI をめぐる政策形成は目まぐるしいスピードで発展し，AI に関する具体的な規制も導入されるようになってきた。そこでは，複雑かつダイナミックな AI 技術の特徴を踏まえ，①リスクベースアプローチ，②ハードローによる大枠の設定とソフトローによる具体化，③プロダクト要件とマネジメント要件の組み合わせ，④共同規制，⑤透明性・アカウンタビリティの重視，⑥認証・監査，⑦継続的な評価と改善，⑧国際的な協調，といったアプローチが共通の要素となっていることを示した。これを本稿では「AI 規制のアプローチ」と呼んでいる。（Ⅱ）

しかし，AI の出力は過去のデータをベースとした統計と確率に基づくことを考えれば，そのリスクは人間によるリスクと相対的であり，AI を固有のリスク源と捉えることには疑問がある。また，そもそも AI と人間は相互に浸潤しながら作用しリスクを生じているのであって，「AI」という技術要素のみに着目して規制を行うことは，リスクの本質を見逃し社会的コストを増大させるおそれがある。さらに，AI 規制のアプローチは，AI 以外の分野の規制において議論されてきたこととも大きく重なるのであり，これを AI という限られた技術の議論に閉じ込めておく必要は見当たらない（Ⅲ）。

既に社会のあらゆる分野において AI の実装が進みつつある現在，AI 規制のアプローチは，業界の別を問わず検討されるべき課題である。また，AI 規制のアプローチの根拠とされている予測不可能性・説明不可能性，バリューチェーン上の主体の多さ，技術革新や普及の速さ，外部からの信頼性判断の難しさ，グローバル化，といった要素は，AI だけでなく，デジタル化やグローバル化に伴う社会システムの複雑化，産業構造の変化，そして不確実性の増大という，より大きな文脈

106) なお，蛇足ではあるが明確化のために，本稿の主張は以下のようなものではないという点も示しておく。

1) 本稿は，AI に関する規制が不要だといっているわけではない。自動運転車にしても，フェイクコンテンツの問題にしても，「AI」という要素は，リスクの要因を構成する一技術的要素にすぎないため，「AI」のみを切り取るのではなく，人間と AI その他のシステムの相互作用と，そこから生じるリスクを総合的に評価して規制を設計すべき，というのが本稿の主張である。

2) 本稿は，ハードローが不要だといっているわけではない。法の支配を実質化するためには，ハードローとソフトローを適切に組み合わせ，その形成及び執行過程にマルチステークホルダーかつアジャイルな手続を導入すべきである，というのが本稿の主張である。

3) 本稿は，日本が EU や米国をはじめとする海外の制度を模倣すべきといっているわけではない。規制の範囲から手法まで，各国の制度には多くの相違点があり，試行錯誤が続いている状況である。また，AI によるリスクに対する社会受容性も，国・地域によって異なる。日本においても，諸国における規制の効果や課題を参照しながら，世界にベストプラクティスを提供できるような取り組みを主体的に行うべきである。

の中で生じている変化である。そのような大きな変化の中で，伝統的な「法の支配」は困難に直面しており，それを乗り越える新たな「法の支配」の形こそが，「AI 規制のアプローチ」に見られるようなマルチステークホルダーかつアジャイル型の規制アプローチである。そうだすれば，「AI 規制のアプローチ」は，今後のあらゆる規制に共通する一般的なアプローチとして捉え直されるべきではないか。このようなアプローチは膨大な情報処理を必要とするため，従来の社会では実装が困難であったが，まさに AI を含む近時の情報処理技術を使うことによって，急速に実現可能性を帯びてきている（Ⅳ）。

このように，AI 規制のアプローチを一般的な規制の議論に昇華させることで，AI 規制の文脈で蓄積された多くのノウハウを他の領域に応用することができるだけでなく，規制の複雑化を避けられ，イノベーションの促進と安全性の確保の両立できるといったメリットがある。他方で，AI 規制論自体が現在進行中でありベストプラクティスが確立されていないこと，専門人材の育成や学際的な協力が必要となること，さらには我々一人ひとりのマインドセットを変えていく必要があることなど，今後の課題も多い（Ⅴ）。

以上の本稿の狙いは，現在 AI 規制論が直面している，課題に溢れた混沌とした状況を，AI という限られたドメインに閉じ込めておくのではなく，法制度や統治全般の問題として法学界，いや社会全体に開放しようというものである。それは，法の支配や民主主義といった社会の根幹を，現代社会に最適な形にアップデートするための困難かつ創造的な航海の第一歩だ。ここまでの長い話にお付き合いいただいた読者の皆様も，ぜひこのブルーオーシャンに身を投じてみませんか？

（はぶか・ひろき）