総務省 学術雑誌『情報通信政策研究』 第 7 巻第 1 号
Journal of Information and Communications Policy Vol.7 No.1

特別寄稿

AI 規正論

新保　史生 [1]（慶應義塾大学）

要　旨

本稿は、①AI システムの研究開発から利用、販売及びサービスの提供にあたって必要な「ルール（規制）」を定め、②その遵守について自主的な取り組みを尊重しつつ、③販売やサービス提供において「事実上の強制規格」として機能する「ルール（整合規格・技術標準・要求事項）」を導入し、④それを計画、実施、評価及び改善するためのマネジメントシステム規格を定め、⑤これらの仕組みを規律するための根拠を法定するとともに、⑥「AI 規正委員会（仮称）」を設置し、⑦「日本版 AI システム適合性評価制度」を中核とする AI 規制構想を提案する。

専ら自主的な規律に期待するソフトローの検討を試行錯誤し続けるのではなく、一方で、反対意見が根強い規制（実質的な禁止事項の法定等）の導入に伴うハードローへの抵抗感を払拭するため、これまで検討がなされてきた原則・指針やガイドライン等をめぐる議論からは発想を転換した取り組みを模索することが本稿の目的である。当該目的を達成するために、規範の遵守を自主性に委ねハードローによる規制を行わない法規制回避論からの脱却、国際的な動向を踏まえた AI 規制の「最適化(optimisation)」、AI の研究開発・利用における将来的な AI 規制政策に資する方策により、新たな AI 規制の制度設計を試みる。

キーワード：AI 規制、整合規格、適合証明、強制規格、マネジメントシステム

１．AI 規制構想の提案意図

１．１．AI 規正論とは

本稿は、AI 時代に向けて必要な AI の安全・安心な利用のための管理及び規制の方策として、新たな「AI 規正論 [2]」を展開することを目的とする。

今後、世の中で広く用いられ日常生活におけるシステムやサービスを制覇することになる AI について、その管理やガバナンス等の規制の主導権を握ることが、今後の AI 時代における AI 政策の覇権を獲得することになる。

---

[1] 慶應義塾大学総合政策学部教授

[2] 法律用語の「きせい」には、「規制」、「規正」、「規整」がある。「規制」とは、「ある事柄を規律し、統制すること。『規正』、『規整』と同様に規律を中心観念とする語であるが、その規律の目的や成果を統制するということに重点を置いて考える場合に多く用いられる。法令上は、『規正』はある事柄を規律して公正な姿に当てはめること、『規整』はある事柄を規律して一定の枠に納め整えることという意味についてのみ、それぞれ用い、それ以外の場合には「規制」を用いることとされている。」『有斐閣法律用語辞典 第５版』（2020）185 頁。本稿は、AI 規制に関する議論や検討事項を規律して公正な AI 規制論の展開を試みるものであるから「規正」の用語を用いることとする。

日常的に利用される AI に求められるのは安全と信頼性が確保され、安心[3]して使うことができることである。それを実現するために、EU は域内の AI 規制の整合化を図るための「整合規則」を「AI 法案」として提案し、同法に基づいて定められる「整合規格(Harmonized Standards[4])」に準拠した高リスク AI に製品安全規制同様の CE マークを付すこと（CE マーキング）で、輸入から販売に至るまでの EU 市場への上市規制を設ける制度の整備を進めている。

我が国にも電気や電子製品の様々な整合規格や JIS 規格・ISO 規格・IEC 規格等の整合標準が存在する。さらに、企業の法令遵守における取り組みにおいても、品質、環境、労働安全衛生などの ISO 規格、JIS Q 15001 を用いたプライバシーマーク、JIS Q 27001（ISO/IEC 27001）を用いた ISMS（情報セキュリティマネジメントシステム）など、様々なマネジメントシステム規格に基づく取り組みが行われ導入実績も豊富である。つまり、EU が AI 法の制定により AI 統治(governance)のための制度として構築を目指している製品安全規制の枠組みは我が国においても導入することが可能であり、その仕組みを「事実上の強制規格（法律に基づく義務）」として日本版の AI 整合規格を策定し適合性評価制度を整備・展開することが可能か検討を試みる。

１．２．日本版の AI システム適合性評価制度の構築を目指す提案の方向性

本稿が示す試案を実際に達成するための選択肢としては、(a)EU が目指している整合規格の整備に向けた取り組みに歩調を合わせ、EU の制度との相互運用性を確保する国内規格をわが国おいても導入する方法（ISO 規格と JIS 規格の関係と考えるとよい）、(b)我が国独自の技術基準として日本版の AI 整合規格を策定し、AI システムの製造・開発、輸入、販売又はサービス提供を行う事業者に当該規格に基づく技術基準適合義務を課し、適合証明の表示を義務付け（PSE マークを想定）、当該適合証明の取得に必要なマネジメントシステム規格を新たに設けることが考えられる。

後者の(b)の選択肢は、いわゆる「ブリュッセル効果[5]」への闘いを挑むものであるが、本稿の提案を端緒に我が国の政策立案の中心地である霞ヶ関から「カスミガセキ効果（仮称）」を今後 AI 関連政策分野で発揮できるかどうかと問われると、既に整合規格の根拠となる

---

[3] 「安心」は、「安全」と「信頼性」の両者が確保されることによって実現するが、AI の信頼性と安全確保と品質保証の観点からの日本国内におけるガイドラインとしては、石油コンビナート等災害防止 3 省連絡会議（経済産業省、総務省消防庁、厚生労働省）「プラント保安分野 AI 信頼性評価ガイドライン第 2 版」(2021 年 3 月)、国立研究開発法人 産業技術総合研究所「機械学習品質マネジメントガイドライン第 1 版」(2020/06/30)、AI プロダクト品質保証コンソーシアム「AI プロダクト品質保証ガイドライン 2022.07 版」<https://www.qa4ai.jp/>がある。

[4] Harmonized Standards は、本稿では「整合規格」と表記する。国際標準化における検討においては、「整合標準」と表記されることも多い。

[5] アニュ・ブラッドフォード(著)、庄司克宏(監修・翻訳)『ブリュッセル効果 EU の覇権戦略: いかに世界を支配しているのか』白水社（2022)。①市場規模、②規制能力、③厳格な規制、④非弾力的対象、⑤不可分性の 5 つの条件を充たす場合に市場における規制力を発揮することができる仕組みのこと。

ISO 規格の検討も進んでいることから現実は厳しいことは承知している。しかし、本稿が示す提案は我が国における AI 規制をめぐる議論や検討過程においてこれまで俎上に載ったことがないものであることから、最初から諦めて(a)の道を選択し EU への追従を図るのではなく、AI 整合規格を我が国独自の基準として制定し、①産業標準化法に基づく新たな「AI マネジメントシステム規格（仮称）」を制定するとともに、②電気用品安全法の改正により PSE マークを AI システム適合証明に活用する方法又は日本版 AI 法の制定による AI 整合規格の根拠となる法整備を実施することにより、「日本版の AI システム適合性評価制度」の構築に向けた提案を行いたい。

１．３．着想の背景

この提案の着想の背景は、EU が AI 法案（AI 整合規則提案）において AI 管理のための制度として構築を目指している製品安全規制に基づく法的枠組みである適合性評価制度が、我が国においても導入が可能ではないかという素朴な問題意識に基づくものである。

適合性評価制度を新興技術規制において用いる着想は、ムーンショット研究開発プロジェクト目標１研究開発プロジェクト「アバターを安全かつ信頼して利用できる社会の実現[6]」の提案にあたり、サイバネティック・アバター（CA）を安全かつ信頼して利用できる CA 基盤を構築するために、CA 操作者の認証技術、CA 認証技術、遠隔操作者が法律に基づいて CA を公的に使用できることを証明・認証する CA 公証に関する仕組み[7]を考えるにあたって、当該認証・公証に係る CA の適合証明（適合性評価制度）のあり方を検討する必要があるとの認識に至ったことに端を発する。

また、汎用性が高いテキスト生成 AI をはじめとする生成系 AI への注目とともに出現している「にわか規制論[8]」に対し、CA のみならず生成 AI などの新興技術の社会実装と社会的受容性を確保するための制度的な課題への応用も可能である。

本稿は、筆者のこれまでの 10 年近いロボット法研究における様々な提案の中でも、今後の新たな展開の端緒になるものであると考えている。初期の提案であった「ロボット法　新８原則（新保試案）[9]」は、自律型ロボットを将来的に社会で受け入れるために共通の認識

---

[6] ムーンショット研究開発プロジェクト　目標１研究開発プロジェクト「アバターを安全かつ信頼して利用できる社会の実現」
<https://www.jst.go.jp/moonshot/program/goal1/15_shimpo.html>。

[7] 新保史生「サイバネティック・アバターの存在証明 ―ロボット・ AI・サイバーフィジカル社会に向けたアバター法の幕開け」人工知能 36 巻（2021）PP.570-577。

[8] 「生成 AI の汎用性に伴う具体的なリスクを見通すことができずこれまでの議論の蓄積が役に立たないかのような錯覚に陥り」、「そのリスクの抽象性ゆえに現在に至るまでの議論がどの程度有用なのか評価できていない」がゆえに唐突な規制論が展開される傾向があることを表すため、新保史生「生成 AI と AI 規制」三田評論 1278 号(2023)43 頁において用いた用語。

[9] 新保史生「何故に『ロボット法』なのか」ロボット法学会設立準備研究会(2015 年 10 月 11 日）報告資料(2015)< http://www.robotlaw.jp/archives/66>。当該原則の詳細については、新保史生「ロボット法をめぐる法領域別課題の鳥瞰」情報法制研究創刊号９章（2017）PP.PP.65-78、Fumio Shimpo, *The Principal Japanese AI and Robot Strategy*

として必要な原則の策定が必要であるとの考えに基づくものであった。2015 年 10 月 11 日に開催した「ロボット法学会設立準備研究会」においてこの案を公表した時点では、あくまで将来的なロボット共生社会に向けて求められる基本となる原則の考案と原則策定の必要性を提案したにすぎない。

その後、AI ブームとともに原則や指針策定の機運が社会的にも高まり、具体的な原則・指針・ガイドライン等が公表されるにつれ、AI 原則を単なる非拘束的な原則として活用を求める段階から、基本法の整備による法令事項としての組み込みや法定公表事項としての位置付けに移行することを検討してよいのではないかとの提案を公表[10]した。しかし、非拘束的な原則の実効性に疑問を呈しつつも、具体的な規制の在り方について方向性を見出すことができない状況が続いていたが、本稿により、ようやく AI 規制の次のステップに向けた方向性を表明する段階に至ったと考えている。

２．日本の AI 関係政策の系譜と法規制回避論の功罪

２．１．AI・ロボット政策の端緒

AI 及びロボットの利用をめぐる基本政策は、「ロボット新戦略（Japan's Robot Strategy ―ビジョン・戦略・アクションプラン）」が 2015 年２月 10 日に日本経済再生本部決定として公表されたことに端を発する。本戦略は AI ブーム前の戦略であるため AI に関する言及はないが、従来の産業用ロボットにとどまらず、ロボットの概念を広く柔軟に捉え、「①世界のロボットイノベーション拠点－ロボット創出力の抜本的強化」、「②世界一のロボット利活用社会」、「③世界をリードするロボット新時代への戦略」をロボット革命の実現に向けた戦略の三本柱としている。

さらに新興技術を活用するために、総合科学技術・イノベーション会議は、「第５期科学技術基本計画」を策定し 2016 年１月 22 日に閣議決定している。基本計画では、Society5.0 の実現のために人工知能技術も重要な役割を担うことに加え、科学技術イノベーションと社会との関係深化の重要性、そのために倫理的・法制度的・社会的取組を行うべきとしている。また、「『超スマート社会』の実現に向けた共通基盤技術や人材の強化」として、AI 等の重点的に取り組むべき技術課題等を明確にし、関係府省の連携をはかり戦略的に研究開発を推進することを明示し、その後、AI・ロボット関係の様々な施策が立案されることとなった。

これらの公表後、日本の政策動向として具体的な検討が進んだ課題としては、自動運転車の公道走行、無人航空機の飛行ルールの整備、AI の研究開発の促進と原則策定の取り組みがあげられる。

---

and Research toward Establishing Basic Principles, RESEARCH HANDBOOK ON THE LAW OF ARTIFICIAL INTELLIGENCE, Woodrow Barfield, Ugo Pagallo（ed）, Edward Elgar Publishing（2018）PP.114-142.

[10] 新保史生「AI 原則は機能するか？ - 非拘束的原則から普遍的原則への道筋 -」情報通信政策研究第３巻第２号（2020）PP.53-70。

２．２．日本政府の AI 戦略

AI に関する政策立案を検討する主な会議として、(1)イノベーション政策強化推進のための有識者会議「AI 戦略」（AI 戦略実行会議[11]）、(2)AI ステアリングコミティー[12]、(3)新 AI 戦略検討会議[13]がある。2023 年５月 11 日には(4)イノベーション政策強化推進のための有識者会議が AI 戦略会議として新たな体制で検討を開始し、(5)AI 戦略チーム（関係省庁連携）も組織されている。

（1）イノベーション政策強化推進のための有識者会議（AI 戦略実行会議）は AI 戦略を立案し、2019 年６月に策定した「AI 戦略 2019」において四つの戦略目標を掲げている。この戦略目標を実現すべく、教育改革、研究開発体制の基盤づくり、社会実装、データ関連基盤整備、AI 時代のデジタル・ガバメント、中小企業・ベンチャー企業の支援、倫理、その他に関する各種取組が推進されている。その後、「AI 戦略 2019」フォローアップとして「AI 戦略 2021」が公表され、、新型コロナウイルス感染症によるパンデミックや地殻変動などより明白になる多くのリスク要因などを反映し、従来のＡＩ戦略の状況に適合した拡張を行った戦略方針として、「AI 戦略 2022」が 2022 年４月 22 日に公表されている。

AI 戦略では、「人間尊重」、「多様性」、「持続可能」の３つの理念のもと、Society 5.0 の実現と SDGs への貢献のため、３つの理念の実装を念頭に５つの戦略目標（人材、産業競争力、技術体系、国際に加え、差し迫った危機への対処）を設定している。特に、AI 戦略 2022 においては、社会実装の充実に向けて新たな目標を設定して推進するとともに、パンデミックや大規模災害等の差し迫った危機への対処のための取組を具体化している。なお、AI に関しては、経済安全保障の観点の取組も日本政府は重要な政策として認識していることを踏まえ、経済安全保障との関係における新興技術の研究開発とその保護への取り組みなど、日本政府として効果的かつ重点的にその施策を進めるため、関係施策の調整や、量子やバイオ等の戦略的取組とのシナジーを追求すべきことを提示している。

（2）AI ステアリングコミティーは、AI 戦略に掲げた基盤的・融合的な研究開発内容（開発工程表作成）、成果の発信、内外のコミュニケーション戦略を策定している。

（3）新 AI 戦略検討会議は、「AI 戦略 2021」を公表した後の新たな政策課題について社会実装に向けた取り組みを推進するため、次の AI 戦略の策定に向けた検討を行うため設置された。この検討会議は、「５年後の利益創出につながる AI の社会実装の促進及び産業競争力の強化」という方針に基づき検討を行っている。最新の国内外の動向を踏まえ、社会実装の充実に向けて新たな目標を設定して推進するとともに、パンデミックや大規模災害といった非日常への対処に係る取組の具体化などを念頭に議論がなされた。さらに、「デジタル社会の実現に向けた重点計画」が 2021 年 12 月 24 日に閣議決定され、AI に関する政策として、データ活用を支える高度コンピューティング技術の研究開発・実証（AI の社会実装に向けた取組の加速）について、深層学習の理論体系や知識融合型 AI 技術、大阪・関西万博での利用を目指す多言語同時通訳等の研究開発を行うことや、AI のブラックボックス

---

11 イノベーション政策強化推進のための有識者会議「AI 戦略」（AI 戦略実行会議）
<https://www8.cao.go.jp/cstp/ai/senryaku/kaigi.html>
12 AI ステアリングコミッティー
<https://www8.cao.go.jp/cstp/ai/steering/steering.html>。
13 新 AI 戦略検討会議<https://www8.cao.go.jp/cstp/ai/shin_ai/shin_ai.html>。

問題解決に向けた説明可能な AI の研究開発を進めることが示されている。

（4）イノベーション政策強化推進のための有識者会議は、新たな体制で AI 戦略会議としての検討を開始し「AI を巡る主な論点」で三つの論点を提示し、（5）AI 戦略チーム（関係省庁連携）は AI 戦略会議における議論等を踏まえ、様々な課題に対して関係省庁連携して迅速に対応する体制を整備し、本稿執筆時点において目下議論の最中である。

２．３．日本の AI 規制に向けた取り組み

AI 戦略全般の推進とともに、日本政府における AI 規制に向けた取り組みは、総務省の AI ネットワーク社会推進会議が 2017 年 7 月に「国際的な議論のための AI 開発ガイドライン(案)」を公表したのをきっかけに、2019 年 8 月には「AI 利活用ガイドライン」、2018 年 6 月には経済産業省が「AI・データの利用に関する契約ガイドライン」、2019 年 3 月には内閣府が「人間中心の AI 社会原則」、2021 年 7 月には経済産業省が「AI 原則実践のためのガバナンス・ガイドライン ver. 1.0」を公表している。2023 年 5 月には、文部科学省「デジタル学習基盤特別委員会」が「生成 AI の学校現場での取扱いに関する今後の対応について」を公表している。

2023 年に開催された G7 広島サミットにおいて、AI ガバナンスに関する国際的議論のための枠組みとして「広島 AI プロセス」を進めることが示され検討がなされている。2023 年 9 月 8 日には、G7 構成国・地域のほか関係国際機関が参加し、生成 AI を巡る国際的なルール形成に向けた議論を行い、成果文書として、「G7 広島 AI プロセス G7 デジタル・技術閣僚声明 [14]」が採択されている。

閣僚声明の主なポイントは、(1)OECD レポートに基づく優先的なリスク、課題、機会の理解（G7 共通の優先的な課題・リスク及び機会を特定）、(2)高度な AI システム（基盤モデルや生成 AI を含む。以下同じ。）に関する国際的な指針（guiding principles）及び行動規範（code of conduct）の策定（AI 開発者を対象とする国際的な行動規範の策定が国際社会の喫緊の課題の 1 つであるという共通認識の下、行動規範策定の基礎として、AI 開発者を対象とする指針の骨子を策定。年内に、開発を含むすべての AI 関係者向けの国際的な指針を策定）、(3)偽情報対策に資する研究の促進等のプロジェクトベースの協力（国際機関と協力し、AI によって生成された偽情報を識別するための最先端の技術的能力に関する研究の促進等、プロジェクトベースの取組を推進することを計画）が示されている。

行動規範策定のための指針の骨子（表 1 参照）では、「国際的に認知された技術標準の開発及び整合性確保の推進」が示されているとともに、各項目の実施は「法的枠組みから自主的なコミットメントその他のさまざまな手段、あるいはそれらの組み合わせ [15]」による実現するとしていることから、本稿が提案する AI 規制の新たな方策が活用される機会を期待したい。

---

14 「広島 AI プロセス閣僚級会合の開催結果」<https://www.soumu.go.jp/menu_news/s-news/01tsushin06_02000277.html>。

15 「G7 広島 AI プロセス G7 デジタル・技術閣僚声明（仮訳）」（2023 年 9 月 7 日）<https://www.soumu.go.jp/main_content/000900471.pdf> 4 頁。

【表１：G7 広島 AI プロセス G7 デジタル・技術閣僚声明における行動規範策定指針骨子】

高度 AI システムの適切な安全対策及び導入前の社会的リスクの考慮
高度 AI システム導入後の脆弱性の特定と低減に向けた努力
モデルの能力、限界、適切・不適切な利用領域の公表
AI 開発者と政府、市民社会、学界との間での責任ある情報共有
プライバシーポリシー及び AI ガバナンスポリシー等のリスク管理計画及び低減手法の開発及び開示
サイバーセキュリティ及びインサイダー脅威対策を含む強固なセキュリティ管理措置への投資
電子透かし技術等の AI が生成したコンテンツを利用者が識別できる仕組みの開発及び導入
社会、環境、安全のリスクを軽減するための研究・投資の優先的な実施
気候危機等の世界最大の課題に対処するための高度な AI システムの優先的な開発
国際的に認知された技術標準の開発及び整合性確保の推進

２．４．法規制回避論の功罪

これまでのところ、AI 規制論は「肯定論」又は「否定論」の両極とともに、規制の必要性を認識しつつ、法規制としていわゆるハードローによる規制ではなく、事業者や民間団体等による純粋な自主的な規律でもなく、政府の検討会が定めたガイドライン等による取り組みを推進する「折衷論」が我が国においては一貫して維持されてきた。しかし、OECD で「AI に関する理事会勧告」(2019 年５月 22 日）[16]が採択された後は、我が国からの提案も踏まえて国際機関において原則策定にまで至った達成感と AI ブームの終息が重なり、次のステップに必要な本来の「規制論」に向けた検討が停滞してしまった。

「AI 原則実践のためのガバナンス・ガイドライン ver. 1.0」も、「D. 今後の課題」において、「(1) 非拘束の中間的なガイドラインを利用するインセンティブの確保」として、「非拘束の中間的なガイドラインは、法的に非拘束であるため、当該ガイドラインを利用するインセンティブが不十分となり、AI 原則を尊重した AI の社会実装の促進という目標を十分に達成できない可能性がある。」と指摘[17]している。

法規制回避論の問題は、原則やガイドラインの策定及び公表段階において意識はされてきたものの、政府における取り組みとしては法的拘束力を前面に打ち出すことが難しいがゆえに、結果的にその妥協案としての非拘束的なガイドライン等に基づくソフトロー路線を維持する折衷論によるしかなかったことはやむを得なかったといえよう。

ゆえに、今後の AI 規制のあり方を考えるにあたっては、我が国の AI 政策における法規制回避論の功罪を省察すべきである。

---

[16] OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449.

[17] 経済産業省：AI 原則の実践の在り方に関する検討会：AI ガバナンス・ガイドライン WG「AI 原則実践のためのガバナンス・ガイドライン ver. 1.0」（令和３年７月９日）31 頁。

法規制回避論の効用としては、我が国同様に自主的な取り組みを尊重しているとされる米国の取り組みと同調することによるメリットがあげられる。一方で、いわゆるブリュッセル効果を発揮しつつある EU の取り組みを傍観することによる AI 規制への乗り遅れの問題がある。現時点における国際的な AI 規制の動向からすると、法規制を回避する方針は功罪相半ばするといった状況である。

ところが、法規制を回避してきたはずの米国は AI 規制に向けた議論に既に着手している。ムーンショット研究開発プロジェクト目標 1「アバターを安全かつ信頼して利用できる社会の実現」に慶應義塾大学大学院政策・メディア研究科訪問教授として参加しているボストン大学ロースクールのウッドロー・ハルツォーグ（Woodrow Hartzog)教授は、2023 年 9 月 12 日に開かれた「米上院プライバシー・テクノロジー・法小委員会 [18]」において、「AI の自主規制は失敗に終わると考えられる(AI self-regulation will be doomed to fail)」との発言とともに、AI の法規制の必要性を証言している。

また、同教授からの情報によると、リチャード・ブルメンタール（Richard Blumenthal）上院議員（民主党）とジョシュ・ホーリー（Josh Hawley）上院議員（共和党）が、AI 規制について法制化に向けた検討を行っており、独立した監督機関の設置、被害に対する法的責任の確保、国家安全保障、透明性の促進、消費者と子供の保護について具体的な規制を行う方針であることや、高度で汎用性が高い AI について、監督機関への登録義務とライセンス制度を導入する方向での検討に着手しているとのことである。

２．５．AI 規制に向けた研究や検討で後塵を拝しつつある要因

我が国における AI 規制に関する議論の遅れを指摘する意見も見受けられるようになりつつあるが、国内における検討や議論は諸外国に先駆けて精緻な取り組みがなされてきたことを再認識すべきである。将来的な AI やロボットの研究開発及び利用における規制に向けて、統一的かつイノベーションの促進に資するために必要な共通認識として、「原則」策定の必要性を提唱し国際的な議論を先行してきたことは明らかである。

にもかかわらず、AI の研究開発だけでなく、AI「規制」に向けた研究や検討においても周回遅れになりつつあるのはなぜか。本稿の当初の原案は国内の AI 関係の文献に基づき今後の AI 規制の方向性を探るための基礎となる研究を行う予定であった。しかし、国内外のガイドラインの紹介、各国の法制度の詳細な調査や国際動向の把握、AI に関する法的諸課題の検討などに関する論考は数多く公表されているにもかかわらず、日本の AI 規制の方向性やそのための戦略を具体的に示している論考が見当たらない。

その要因として考えられる問題を文献整理により若干把握できたので記しておきたい。なお、以下の懸念事項を示す根拠として該当する文献を明記すべきであることは当然認識しているが、相当批判的な視点からの懸念事項であるため特定の文献明示は行わないこととする。

①規制・禁止同視論

[18] U.S. Senate Committee on the Subcommittee on Privacy, Technology, and the Law, *Oversight of A.I.: Legislating on Artificial Intelligence*, <https://www.judiciary.senate.gov/committee-activity/hearings/oversight-of-ai-legislating-on-artificial-intelligence>(Tuesday, September 12th, 2023).

「規制」を「禁止」と同義に議論がなされる傾向があること。イノベーションの促進への仮想・架空の懸念（実際には阻害されるイノベーションそのものが存在していない）。緩和する規制が存在しないにもかかわらず規制緩和を主張する見解など。

②規制の不存在の反射としての躊躇

規制されていないので「実施できない」という不合理な理屈（本来は規制されていないので実施して何ら問題ないはず）。新たな技術開発やサービスの提供にあたっての「規制の不存在」による事業展開や利用への躊躇。

③責任転嫁論

技術開発や利用の遅れを法整備の不備（規制の不存在としての）を理由に言い訳をしているとしか思えない指摘。

④法と倫理の混同

法と倫理を区別せずAIをめぐる法的課題を一緒くたに倫理的課題にしてしまい、情報倫理で情報法に関する課題をすべて論じようとしていた時代を彷彿させるもの。

⑤取り越し苦労的な懸念先行論

まずは懸念やリスクに関する問題を先行して議論し、振り返ってみるといずれの懸念も単なる杞憂と帰してしまっている指摘。リスク「認識」ではなく単なるリスク「例示」による自己満足的な議論により本来のリスク認識が達成できていない。抽象的で中途半端なディストピア的議論。AIについての抽象的畏怖に基づく浅薄なリスク論、AIによる人類駆逐懸念やラッダイト運動的排斥主張など。

⑥クリックベイト的論文

AIによる問題に関する議論と謳っていながら、その内実は単なる情報システムの高度化への懸念論でしかない論考（DX推進を謳いながら実際には基幹系刷新にすぎず現行システムの再構成を売り込んでいるようなもの）。タイトルにAIと冠しているためAIに関する論考かと思いきや無関係の内容で、まるでクリックベイトのような欺瞞的なものや、AとIという文字列から成る空想の産物に関する内容であったり、電子法人格(electronic person)に関する議論がいつの間にかAI基本的人権論に飛躍しているものもある。

⑦検討事項の断片的抽出・認識・評価による弊害

単に思いついた課題から議論をはじめる断片的な検討事項の抽出・認識を行う指摘も弊害が大きい。例えば、新たな技術の利用に伴い議論される法的課題として、まずは、知的財産やプライバシー関係の問題からとりあえず議論をはじめる傾向があること。その結果、真に検討すべき論点の欠落が生じ中途半端な楽観論により議論が収束してしまうことがある。なぜその原則が必要なのか理由を説明できないにもかかわらず、単に流行に遅れんとせんがために、適当にピックアップした原則を提示しているだけの指針。AI・ロボットに関する法的課題を産業用ロボットの延長でしか考えていない場合もある。

以上のようなAI規制をめぐる研究や議論の仕方における問題を意識し、建設的なAI規制に向けた議論を進めることも必要ではないか。

２．６．歴史は繰り返す

1980年代以降に進展したデータのコンピュータ処理、1990年代のインターネットの登場、携帯端末やスマートフォンの普及、2000年代以降のSNSなどユーザが情報を生成す

る環境の定着、そして、サイバー空間（仮想空間）とフィジカル空間（現実空間）が別個の空間として存在してきた時代から、両者が高度に融合し併存する「サイバー・フィジカル時代」が到来しようとしている。データの利用やネットワークにおけるサービスが充実するにつれ、新たなサービスに対する規制や利用者保護のあり方が問われてきたが、この過程において繰り返されてきたのは、法規制をはじめとする新たな規制の試みは、イノベーションの阻害になるといったような定型的な批判や包括的な規制に反対であるといった各方面からの意見である。これにより本来あるべき規制が実現できないことが多々あり、法規制とイノベーションの促進を両立するための視点が欠けているがゆえに、結果的に我が国の成長の阻害要因になっている面があることは否定できない。

これまでの過去の例を若干振り返ってみても、2003 年に成立した個人情報保護法案が審議されたときは、マスコミ規制法であるとの批判とともに個人情報の取扱いについて事業の支障になるという指摘がなされた。EU の「一般データ保護規則（GDPR)）は厳しい規制であるがため日本はその取り組みに反対すべきであり、そのような厳しい規制はデータ利活用の観点から不要であるという見解も表明されてきた。EU の AI 規制への日本国内の反応においても、包括的な AI 規制は時期尚早であるとか産業界の負担といった懸念が示されており [19]、EU の AI 法案への対応（批判）もその二の舞を演じるかの様相を呈している。

３．EU の AI 規制法体系

３．１．EU の AI 法案（整合規則提案）、機械規則提案、AI 法的責任指令案、データ法案

EU の将来的な AI 規制法体系は、①AI 法案（整合規則提案）[20]、②機械規則提案 [21]、③AI 法的責任指令案 [22]及び④データ法案 [23]から構成される。

---

[19] 日本経済新聞「EU の AI 規制案、リスク 4 段階に分類　産業界は負担増警戒」2021 年 4 月 22 日。

[20] European Commission, Proposal for a Regulation of The European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM(2021) 206 final, 2021/0106 (COD)Brussels,21.4.2021.

[21] European Commission, Proposal for a Regulation of the European Parliament and of the Council on machinery products, 21.4.2021 COM(2021) 202 final2021/0105 (COD). 機械指令(Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast),OJ L 157, 9.6.2006, p. 24–86)が機械規則に改正される。

[22] Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)(COM/2022/496 final), Liability Rules for Artificial Intelligence<https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en>, 夏井高人「AI 法的責任指令案」法と情報雑誌第 8 巻第 1 号（通巻第 51 号・2023 年 9 月）PP.1-80。

[23] European Data Governance Act<https://digital-strategy.ec.europa.eu/en/policies/data-governance-act>.

本稿が示す構想は、EU の AI 法案が定める整合規格が事実上の強制規格として今後機能すると考えられる点に着目したことが着想の端緒である。したがって、日本版の AI 整合規格や適合性評価制度の提案に向けた検討を行うにあたっては、EU の AI 規制法体系及び AI 法案が定める高リスク AI に適用される「適合性評価」の仕組みを理解することが不可欠である。

EU における AI 規制の枠組み[24]は、①EU 市場に投入され利用される AI システムの安全規制を、基本的権利と EU の価値を保護する既存の法令に基づき実施すること、②AI への投資とイノベーション促進、③基本的権利の保障と安全性確保のため AI システムへのガバナンスと効果的な法執行、④信頼できる AI により単一市場の発展を促進し市場の断片化を防ぐことにある。新たな法的枠組みは、既存の法令に基づく加盟国レベルでのガバナンスシステムと、欧州 AI 委員会（European Artificial Intelligence Board: EAIB）の設置による EU レベルでの協力の双方の取り組みによって実現される。

３．２．EU の AI 法案の名称について

2021 年 4 月 21 日に公表された AI 法案の名称の原題文は、Proposal for a Regulation of The European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts である。2023 年 6 月 14 日には欧州議会において、Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts[25]が採択されている。

筆者は、欧州委員会が提出した規則提案については、「人工知能に関する整合規則（人工知能法）の制定及び関係法令の改正に関する欧州議会及び理事会の規則提案」（略称：AI 整合規則提案）と訳出し、欧州議会採択案は、「欧州議会が 2023 年 6 月 14 日に採択した人工知能に関する整合規則（人工知能法）の制定及び域内の関連法令の改正に関する欧州議会及び理事会の規則提案に関する修正案」と訳出しているが、「人工知能に関する調和の取れ

---

[24] European Commission, A European approach to Artificial intelligence <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>.

[25] Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))<https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai>, <https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html>.

たルールを定める規則の提案」と訳している場合[26]も多い。

AI 法案が「整合規格」を核とする「整合規則」の法定を目指しているという趣旨を理解し、整合規則の目的と意図を把握する上で、AI 法案の条文訳については、夏井高人「人工知能に関する整合化された規定を定め、欧州連合の一定の立法行為を改正する欧州議会及び理事会の規則（人工知能法）の提案（COM/2021/206 final）」法と情報雑誌 6 巻 5 号（2021 年 11 月）を参照しその趣旨について適切に理解することが必要であると考えられる。

３．３．AI 法案の目的

AI 法案は、AI システムをそのリスク[27]に応じて、①受容できないリスク、②高リスク、③限定的なリスク、④低リスク又はリスク無しの四段階のリスクに分類。当該リスク分類に応じて、①利用禁止 AI、②高リスク AI、③特定の AI、④低リスク・無リスク AI に分けて規定している。

AI 法案の目的は、「従来から EU 市場に上市する製品の製造者や輸入者等に課されている製品安全規制同様の義務を高リスクに分類される AI システムにも拡充して CE マーキングの対象とし、そのための適合性評価および第三者認証制度について定め、新たな整合法令の整備を目指すもの[28]」である。

３．４．AI 法案の構成

AI 法案の条文のうち、リスクベースに基づく規制のアプローチと適合性評価をはじめとする AI 法案の主たる規制内容を整理すると以下の通りである。

---

[26] AI ネットワーク社会推進会議「報告書 2022　～「安心・安全で信頼性のある AI の社会実装」の更なる推進 ～」2022 年 7 月 25 日 2 頁。

[27] European Commission, Regulatory framework proposal on Artificial Intelligence, 26 April 2021 <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>, Excellence and trust in artificial intelligence<https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_en>.

[28] 新保史生「EU 新 AI 整合規則提案にみる AI 規制戦略の構造・意図とブリュッセル効果の威力」ビジネス法務 2021 年 8 月号(2021)188 頁。新保史生「EU の AI 法（AI 整合規則提案）の制定に向けた検討とその影響について」日本経済研究センター欧州研究プロジェクト報告書（2023）61-81 頁。

【表２：AI 法案の構成】

（1）利用規制
人工知能の利用禁止行為（５条）
（2）高リスクAIに関する義務
高リスクAIの分類及び対象リスト（６条・７条）
高リスクAIシステムのマネジメントシステム要求事項（８-15条）
高リスクAIシステムのプロバイダ（生成AIを含む）及びデプロイヤ（実装者）等の義務（16-29条）
（3）適合性評価
第三者認証機関（30-39条）
適合性評価等（40-51条）
（4）特定のAIシステムに対する透明性（52条）
（5）行動規範の策定（69条）
（6）イノベーション促進施策
AI規制サンドボックス、小規模事業者・ユーザ支援等（53-55条）
（7）ガバナンス
欧州人工知能委員会（EAIB）の設置等（56-59条）
（8）監督及び法執行
高リスクAIに関するデータベース（60条）
市販後のモニタリング（61条）
インシデント報告義務（62条）
法執行（63-68条）
罰則等（71-72条）

３．５．高リスク AI

「高リスク AI」は、附属書 II[29]に記載されている製品安全規制の対象となる AI システムであり、かつ、(a)安全構成要素において用いられる AI システム、(b)安全構成要素として用いられる AI システムを含む製品として第三者適合性評価を受ける義務があるものの両者を満たす場合をいう。（①機械、②玩具、③海洋レクリエーション船舶、④リフト、⑤爆発性雰囲気装置、⑥無線機器、⑦圧力機器、⑧索道設備、⑨個人用保護具、⑩ガス燃焼機器、

[29] ①機械、②玩具、③海洋レクリエーション船舶、④リフト、⑤爆発性雰囲気装置、⑥無線機器、⑦圧力機器、⑧索道設備、⑨個人用保護具、⑩ガス燃焼機器、⑪医療機器、⑫体外診断用医療機器。ANNEXES to the Proposal for a European Commission, Regulation of The European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM(2021) 206 final, ANNEXES 1 to 9.

⑪医療機器、⑫体外診断用医療機器：６条１項[30]）ただし、２条２項により附属書 II の B[31]（航空機、車両、鉄道、船舶等）（①民間航空、②マイクロカー、③農業・林業用トラクター、④船舶用機器、⑤鉄道システム、⑥自動車及びトレーラー等、⑦無人航空機）は適用外となっており、84 条の評価・見直し条項が適用される。

附属書 II の B のリストに掲げられている AI システムこそ、自動運転、無人の農業用機器、ドローンをはじめ無人の航空(UAV)、車両(UGV)、船舶(USV)など、今後、AI を用いた展開が最も期待される分野である。そのため、規則提案の公表時点では高リスク AI に課す義務の適用については産業界との協議を念頭に留保規定として評価見直し条項の対象としている。

さらに、６条１項が規定する高リスク AI システムに加えて、附属書 III[32]（①自然人の生体識別及び分類、②重要インフラの管理・運用、③教育及び職業訓練、④雇用、労働者管理、自営業へのアクセス、⑤必要不可欠な民間サービスや公共サービス、⑥法執行、⑦移民、亡命、国境管理、⑧司法行政及び民主主義プロセス。）に係る分野における AI システムの利用も高リスクとしている。

高リスク AI のうち、①自然人の生体識別及び分類に係る「遠隔生体識別」は、公的機関による利用は、AI 法案５条に基づき AI の利用禁止行為の対象として禁止されるが、それ以外の利用は認められる。公的機関による社会的スコアリングも利用禁止対象であるが、④雇用、労働者管理、自営業へのアクセスに係る利用は制限されていない。つまり、企業の採用活動や金融機関における与信判断において利用する場合は禁止対象とはならない。

附属書 III の分野は、高リスク AI システムがもたらす危害または悪影響のリスクに応じて 7 条では欧州委員会が見直しを行う権限について定めている。

---

30 ６条１項の対象（附属書 II の A）。1. 機械指令（2006/42/EC 指令）、2. 玩具安全指令（2009/48/EC 指令）、3. 海洋レクリエーション船舶指令（2013/53/EU 指令）、4. リフト指令（2014/33/EU 指令）、5. 爆発性雰囲気装置及び保護システム指令（2014/34/EU 指令）、6. 無線機器指令（2014/53/EU 指令）、7. 圧力機器指令（2014/68/EU 指令）、8. 索道設備規則（(EU) 2016/424 規則）、9. 個人用保護具規則（(EU) 2016/425 規則）、10. ガス燃焼機器規則（(EU) 2016/426 規則）、11. 医療機器規則（(EU) 2017/745 規則）、12. 体外診断用医療機器規則（ (EU) 2017/746 規則）。

31 ６条２項の対象（附属書 II の B）。1. 民間航空安全規則（(EC) 300/2008 規則）、2. 二輪・三輪及び四輪マイクロカー規則（(EU) No 168/2013 規則）、3. 農業用及び林業用トラクター規則（(EU) No 167/2013 規則）、4. 船舶用機器指令（2014/90/EU 指令）、5. 鉄道システム相互運用性指令（ (EU) 2016/797 指令）、6. 自動車及びトレーラー等システム規則（(EU) 2018/858 規則）、自動車及びトレーラー等型式認証規則（(EU) 2019/2144 規則）、7. 無人航空機規則（(EU) 2018/1139 規則）。

32 1.自然人の生体識別及び分類、2.重要インフラの管理・運用、3.教育及び職業訓練、4.雇用、労働者管理、自営業へのアクセス、5.必要不可欠な民間サービスや公共サービス、6.法執行、7.移民、亡命、国境管理、8.司法行政及び民主主義プロセス。

３．６．適合性評価制度と CE マーク

AI 法案に基づく AI システム適合性評価制度は、AI を利用する際に安全についてリスクが高いと考えられる「高リスク AI システム」を EU 市場に上市及び実装（利用可能状態）にするにあたって遵守しなければならない手続的義務であり、①適合性評価、②適合宣言書への署名、③CE マーキングの実施から構成される。事後的な監督手続として、①AI データベースへの登録、②市販後のモニタリング、③インシデント報告義務等が課される。

AI 法案の目的の根幹にあるのは、「AI システム適合性評価制度」の導入を目指すための新たな「整合規則」の制定を目指すものであって、高リスク AI システムの適合性評価及び第三者認証制度の構築にあると評価できる。規則の提案とともに、EU 機械指令（ロボット、芝刈り機、3D プリンター、建設機械、工業用生産ラインなど、消費者向け製品から専門家向け製品まで幅広い分野が含まれる機械製品の利用における健康及び安全要件を定義）も新たな機械規則へと改正される点からもその意図は明らかである。

なお、「新機械規則」は、新世代の機械がユーザや消費者の安全を保証し、イノベーションを促進することを目的[33]としているが、規則提案とともに改正が必要な理由は、EU 域内で販売する指定製品に貼付が義務付けられている製品安全マークである「CE マーク」の表示義務としての「CE マーキング」に係る法的義務を AI システムにも課すことにある。

３．７．適合性評価の構造

サービスや商品を提供するにあたって、それらの品質が特定の基準に達しているか否かを第三者が判断する仕組みは、サービスや商品を安全かつ安心して利用する際に重要である。AI 法案では、その役割を担う組織を「第三者認証機関（Notified Body（以下、「NB」という。）」（30-39 条）として定めている。NB の設置は、各加盟国が適合性評価のための第三者認証機関を設置することが定められ、認証機関の要件等が規定されている。

高リスク AI システムであって第三者適合性評価を受ける義務があるものは、適合性評価等（40-51 条）に関する手続に基づき、高リスク AI システムのマネジメントシステム要求事項（８-15 条)を満たしていることを、NB による審査を受けなければならない。

要求事項は、①リスクマネジメントシステムの構築、②適切なデータガバナンス、③技術文書、④記録保持、⑤透明性及び利用者への情報提供、⑥人的監視、⑦正確性、堅牢性及びサイバーセキュリティ要件から構成される。

第 43 条の適合性評価手続に基づく承認後は、表示する CE マークに認証を受けた NB の識別番号を記載する仕組みとなっており、一般的な第三者認証によるマーク制度同様の手続となっている。

適合性の評価として、整合規格が存在する場合は当該整合規格を満たしていれば要件適合を判断し既に審査を受けている安全規格の範囲内での利用が認められる。（40 条）

一方、整合規格不存在の場合は、欧州委員会が共通仕様を策定し、共通仕様を満たしてい

---

[33] European Commission, Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence, Press release21 April 2021Brussels <https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682>.

れば要件適合を推定する仕組みとなっている。（41 条）

附属書 III の遠隔生体識別システムの適合性評価（44 条）については、整合規格又は共通仕様を「適用する」場合は、NB の関与のもと、①内部統制、②品質マネジメントシステムの評価、③技術文書の評価に基づく適合性評価手順を実施する手続が定められている。

一方、整合規格又は共通仕様を「適用しない」場合は、NB の関与のもと、①品質マネジメント、②技術文書に基づく評価を実施することとなっている。

附属書 III の遠隔生体識別システム以外は内部統制の評価、附属書 II の A 記載の場合は、各法令（各規則や指令）に基づく適合性評価の実施が定められている。

３．８．EU の AI 規制が目指す「共通の強行的な要件（common mandatory requirements）」の効用

本稿が想定する「事実上の強制規格」は、AI 法案において、「整合化された技術標準を通じて更に運用可能なものとなる市場にそれらのシステムが置かれる前に、一定の AI システムの設計及び開発に適用可能な共通の強行的な要件を定義する。提案は、事後の監督が行われる方法を整合化することによって AI システムが市場に置かれた後の状況にも対処している。[34]」とし、「共通の強行的な要件（common mandatory requirements）」としているものである。

AI の研究開発でたとえ優位に立てない場合であっても、適合性評価制度によって市場への投入における参入規制を設けることで、製造販売やサービス提供規制における主導権を握り市場管理を行うことが可能となる。

AI 法案は、リスクベースアプローチや利用禁止対象の AI システムに関する規定が注目されているが、それらの議論に目を奪われているうちに、AI システムに対する適合性評価の仕組みが整備され気が付いた時にはその制度が定める手続きを遵守しなければ、開発した製品を EU 市場で販売することができなくなる。当該規制手法の真髄は、適合性評価を行うことで EU 市場への AI システムを利用した製品やサービスの販売を規制することにあるため、包括的な規制であるとして新たな規制に反対してリスク評価に基づく分類に納得がいかなくても、上市規制が導入された場合にはその手続きに従うほかない。

４．EU における整合規格策定に向けた取り組み

AI 法案の欧州議会採択案では、第 40 条で新たに「欧州標準化規則 [35]」10 条に基づく標

---

[34] 夏井高人「人工知能に関する整合化された規定を定め、欧州連合の一定の立法行為を改正する欧州議会及び理事会の規則（人工知能法）の提案（COM/2021/206 final)」法と情報雑誌 6 巻 5 号（2021 年 11 月）350 頁。

[35] Regulation (EU) No 1025/2012 of the European Parliament and of the Council of 25 October 2012 on European standardisation, amending Council Directives 89/686/EEC and 93/15/EEC and Directives 94/9/EC, 94/25/EC, 95/16/EC, 97/23/EC, 98/34/EC, 2004/22/EC, 2007/23/EC, 2009/23/EC and 2009/105/EC of the European Parliament and of the Council and repealing Council Decision 87/95/EEC and Decision No 1673/2006/EC of the European Parliament and of the Council Text with EEA relevance,

準化要請に関する規定が追加されている。当該規定に基づき、欧州委員会は AI 法案が定める適合性評価制度の実施に必要な整合規格の策定に向けて必要な要求事項の標準化に向けた取り組みに着手している。本稿執筆時点では未だドラフト段階ではあるが、この標準化要請に基づいて今後策定される整合規格が EU の適合性評価制度の中核となると考えられることから、現時点で確認できる事項を紹介する。

なお、本要請案は AI 法案の制定後に発効するため、標準化要請が具体化するのも AI 法制定後となる。

４．１．AI 整合規格の標準化要請案（草案）の名称、根拠及び目的

本草案の名称は、「安全で信頼できる AI を支援するための欧州標準化機関への標準化要請案（草案）[36]」である。

標準化要請の根拠は、AI 法案（整合規則提案）であり、高リスク AI の適合性評価制度を構築するため、欧州標準化規則第 12 条に基づいて 10 分野の整合規格を策定するために必要な標準化要求を決定することが目的である。

安全で信頼できる AI システムを支援するために、欧州規格又は欧州標準化規格類の策定を要請する意図は、「欧州標準化のための 2022 年次統合作業計画」に関する欧州委員会通知 C(2022) 546 の附属文書内の「標準化戦略に関して欧州委員会が設定した標準化緊急課題」と題する表 63 番目の項目記載事項に基づくものとなっている。

整合規格を策定する実施機関は、欧州標準化委員会（CEN）及び欧州電気標準化委員会（CENELEC）が指定されている。なお、欧州の標準化機関の一つである欧州電気通信標準化機構（ETSI）については、セキュリティなど特定の事項に関して既に本文書通知前の段階において作業に着手しているとともに、特定の専門知識を有していることから、作業計画の作成中に CEN および CENELEC が ETSI と協議し、これらの事項に関して ETSI の貢献を認めるための方法を確認することが適切であるとしている。

４．２．国際標準化の効果

AI 関係の国際標準については、ISO/IEC JTC 1/SC 42 において検討がなされ、公表済みの AI 標準は本稿執筆時点（2023 年 8 月）で 20 件[37]策定されている。適合性評価との関係

OJ L 316, 14.11.2012,

[36] European Commission, Draft standardisation request to the European Standardisation Organisations in support of safe and trustworthy artificial intelligence, A Notification under Article 12 of Regulation (EU) No 1025/20121 (5.12.2022).

[37] ① ISO/IEC TS 4213:2022(Information technology − Artificial intelligence − Assessment of machine learning classification performance), ② ISO/IEC 8183:2023(Information technology − Artificial intelligence − Data life cycle framework), ③ ISO/IEC 20546:2019(Information technology − Big data − Overview and vocabulary), ④ISO/IEC TR 20547-1:2020(Information technology − Big data reference architecture − Part 1: Framework and application process), ⑤ISO/IEC TR 20547-2:2018 (Information technology − Big data reference architecture − Part 2: Use

において参照すべきものとして、ユースケース [38]、ライフサイクル [39]、データ品質 [40]、バイアス [41]、ML 分類性能評価 [42]、堅牢性 [43]などがある。

標準化要請案（草案）においても国際標準化の効果について言及しており、信頼できる AI（trustowrthy AI）という共通のビジョンを世界中に定着させるのに役立ち、他方では、AI を搭載した製品やサービスに関連して起こりうる技術的障壁を取り除き貿易を促進することができるとしている。

そのために、規則(EU) No 1025/2012（欧州標準化規則）の第 10 条 1 項に基づく要求事

---

cases and derived requirements), ⑥ISO/IEC 20547-3:2020 (Information technology − Big data reference architecture − Part 3: Reference architecture), ⑦ISO/IEC TR 20547-5:2018 (Information technology − Big data reference architecture − Part 5: Standards roadmap), ⑧ISO/IEC 22989:2022 (Information technology − Artificial intelligence − Artificial intelligence concepts and terminology), ⑨ISO/IEC 23053:2022(Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML))、⑩ISO/IEC 23894:2023(Information technology − Artificial intelligence − Guidance on risk management)、⑪ISO/IEC TR 24027:2021(Information technology − Artificial intelligence (AI) − Bias in AI systems and AI aided decision making)、⑫ISO/IEC TR 24028:2020 (Information technology − Artificial intelligence − Overview of trustworthiness in artificial intelligence)、⑬ISO/IEC TR 24029-1:2021(Artificial Intelligence (AI) − Assessment of the robustness of neural networks − Part 1: Overview)、⑭ ISO/IEC 24029-2:2023(Artificial intelligence (AI) − Assessment of the robustness of neural networks − Part 2: Methodology for the use of formal methods)、⑮ISO/IEC TR 24030:2021(Information technology − Artificial intelligence (AI) − Use cases)、⑯ISO/IEC TR 24368:2022(Information technology − Artificial intelligence − Overview of ethical and societal concerns)、⑰ISO/IEC TR 24372:2021(Information technology − Artificial intelligence (AI) − Overview of computational approaches for AI systems)、⑱ISO/IEC 24668:2022(Information technology − Artificial intelligence − Process management framework for big data analytics)、⑲ISO/IEC 25059:2023(Software engineering − Systems and software Quality Requirements and Evaluation (SQuaRE) − Quality model for AI systems)、⑳ ISO/IEC 38507:2022(Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations).

[38] ISO/IEC TR 24030:2021(Information technology — Artificial intelligence (AI) — Use cases)<https://www.iso.org/standard/77610.html>.

[39] ISO/IEC FDIS 5338(Information technology — Artificial intelligence — AI system life cycle processes)<https://www.iso.org/standard/81118.html>.

[40] ISO/IEC DIS 5259-2(Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 2: Data quality measures)<https://www.iso.org/standard/81860.html>.

[41] ISO/IEC TR 24027:2021(Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making)<https://www.iso.org/standard/77607.html>.

[42] ISO/IEC TS 4213:2022(Information technology — Artificial intelligence — Assessment of machine learning classification performance)<https://www.iso.org/standard/79799.html>.

[43] ISO/IEC TR 24029-1:2021(Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview)<https://www.iso.org/standard/77609.html>.

項を定め、本要請に基づく欧州規格及び欧州標準化規格類[44]の起草にあたっての確認項目として 10 項目からなる要求事項を定めることを目的としている。

４．３．標準化要請

標準化要請案（草案）第１条は、「欧州標準化委員会（CEN）及び欧州電気標準化委員会（CENELEC）に対し、安全で信頼できる AI を支援するため、附属書 I の表１に記載された新規の欧州規格又は欧州標準化規格類を起草するよう要請する。」と定め、「第１項に基づいて策定される欧州規格又は欧州標準化規格類は、附属書 II が規定する要求事項を満たさなければならない。」としている。

なお、本稿執筆時点（2023 年 8 月）で国際標準策定に向けて検討が進んでいる規格は 32 件[45]である。

---

[44] 欧州標準化規則 第２条（定義）では、「規格」、「国家規格」、「EN 規格」、「整合規格」及び「国際規格」が定義されているが、本草案が策定を要求する「欧州標準化デリバラブル（Europeanstandardisation deliverable）」とは、「反復または継続的に適用するために欧州標準化団体によって採択され、かつ、遵守が義務ではない、欧州規格以外の様々な技術仕様を意味する。」（鈴木俊吾、国松麻季「欧州標準化規則（1025/2012）及び Mandate（標準化要求）に係る動向について」国際ビジネス研究第 10 巻第１号（2018）28 頁）とされている。この点につき本稿では、European standardisation deliverable は、「欧州標準化規格類」と訳す。その他の用語については、「ISO/IEC 専門業務用指針，第１部 統合版 ISO 補足指針- ISO 専用手順 第 10 版，2019（Consolidated ISO Supplement - Procedure specific to ISOTenth edition, 2019[Based on the fifteen edition (2019) of the ISO/IEC Directives, Part 1]ISO/IEC 専門業務用指針，第１部，第 15 版（2019）をベースとする英和対訳版」ISO/IEC 原本発行：2019-05-01、英和対訳版発行：2019-07-01 に基づく訳語とする。

[45] ①ISO/IEC DIS 5259-1(Artificial intelligence − Data quality for analytics and machine learning (ML) − Part 1: Overview, terminology, and examples), ②ISO/IEC DIS 5259-2(Artificial intelligence − Data quality for analytics and machine learning (ML) − Part 2: Data quality measures), ③ISO/IEC DIS 5259-3(Artificial intelligence − Data quality for analytics and machine learning (ML) − Part 3: Data quality management requirements and guidelines), ④ISO/IEC DIS 5259-4(Artificial intelligence − Data quality for analytics and machine learning (ML) − Part 4: Data quality process framework), ⑤ISO/IEC CD 5259-5(Artificial intelligence − Data quality for analytics and machine learning (ML) − Part 5: Data quality governance), ⑥ISO/IEC CD TR 5259-6(Artificial intelligence − Data quality for analytics and machine learning (ML) − Part 6: Visualization framework for data quality), ⑦ISO/IEC FDIS 5338(Information technology − Artificial intelligence − AI system life cycle processes), ⑧ISO/IEC FDIS 5339(Information technology − Artificial intelligence − Guidance for AI applications), ⑨ISO/IEC DIS 5392(Information technology − Artificial intelligence − Reference architecture of knowledge engineering), ⑩ISO/IEC DTR 5469(Artificial intelligence − Functional safety and AI systems), ⑪ISO/IEC CD TS 6254(Information technology − Artificial intelligence − Objectives and approaches for explainability of ML models and AI systems), ⑫ISO/IEC

４．４．作業計画

標準化要請案（草案）第２条１項は、「CEN 及び CENELEC は，附属書 I が記載するすべての規格、担当 TB 及び要請された標準化活動の実施期限を示す作業計画を作成しなければならない。」とし、「EU の中小企業及び市民社会組織の効果的な参加を確保し、基本的権利の分野で関連する専門知識を収集するために実施される行動を含む。附属文書第１項の作業計画の作成にあたっては、CEN 及び CENELEC は、欧州電気通信標準化機構（ETSI）と協議し，以下の要素に対する ETSI の貢献を確保するための方法を検討し合意するものとする。」と定めている。

CD TS 8200(Information technology − Artificial intelligence − Controllability of automated artificial intelligence systems), ⑬ISO/IEC DTS 12791(Information technology − Artificial intelligence − Treatment of unwanted bias in classification and regression machine learning tasks), ⑭ISO/IEC CD 12792(Information technology − Artificial intelligence − Transparency taxonomy of AI systems), ⑮ISO/IEC AWI TS 17847(Information technology − Artificial intelligence − Verification and validation analysis of AI systems), ⑯ISO/IEC CD TR 17903(Information technology − Artificial intelligence − Overview of machine learning computing devices), ⑰ISO/IEC AWI TR 18988(Artificial intelligence − Application of AI technologies in health informatics), ⑱ ISO/IEC AWI TR 20226(Information technology − Artificial intelligence − Environmental sustainability aspects of AI systems), ⑲ISO/IEC AWI TR 21221(Information technology – Artificial intelligence – Beneficial AI systems), ⑳ ISO/IEC AWI TS 22440(Artificial intelligence — Functional safety and AI systems — Requirements), ㉑ISO/IEC AWI TS 22443(Information technology — Artificial intelligence — Guidance on addressing societal concerns and ethical considerations), ㉒ISO/IEC AWI 24029-3(Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 3: Methodology for the use of statistical methods), ㉓ISO/IEC CD TR 24030(Information technology — Artificial intelligence (AI) — Use cases), ㉔ ISO/IEC DTS 25058(Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Guidance for quality evaluation of artificial intelligence (AI) systems), ㉕ISO/IEC AWI TS 29119-11(Software and systems engineering — Software testing — Part 11: Testing of AI systems), ㉖ISO/IEC FDIS 42001(Information technology — Artificial intelligence — Management system), ㉗ ISO/IEC CD 42005(Information technology — Artificial intelligence — AI system impact assessment), ㉘ISO/IEC DIS 42006(Information technology — Artificial intelligence — Requirements for bodies providing audit and certification of artificial intelligence management systems), ㉙ISO/IEC AWI 42102(Information technology — Artificial intelligence — Taxonomy of AI system methods and capabilities), ㉚ISO/IEC AWI TR 42103(Information technology — Artificial intelligence — Overview of synthetic data in the context of AI systems),㉛ISO/IEC AWI 42105(Information technology — Artificial intelligence — Guidance for human oversight of AI systems),㉜ ISO/IEC AWI TR 42106(Information technology — Artificial intelligence — Overview of differentiated benchmarking of AI system quality characteristics).

【表３：作業計画で定める実施事項】

(a) 附属書Iの表１の８に基づく欧州規格及び欧州標準化規格類の策定
(b) 附属書 I の表１ において列挙されている欧州規格及び欧州標準化規格類のうち、同表の８に言及されているもの以外のものについては、セキュリティ評価及び統合を実施すること
(c) 附属書 II の第１章に基づくETSIが実施可能な手続及び仕様の策定及び精緻化

草案第２条２項はETSIが実施すべき事項について定めており、「ETSIが実施可能な貢献の方法及びその範囲に関する記述は、第１項において示されている作業計画の下に記載されなければならない。CEN、CENELEC及びETSIが、前項で定めるETSIが実施可能な貢献について合意できない場合、作業計画にはその理由を記載しなければならない。ETSIが実施可能な貢献は、第１条に基づく要請の実行に対するCEN及びCENELECの責任並びに附属書 II に規定された欧州規格および欧州標準化規格類に対する要求事項を損なうものであってはならない。」としている。

草案第２条３項はCEN及びCENELECが実施すべき事項について定めており、「CEN及びCENELECは，第１条に基づく標準化要請が欧州委員会に採択されてから４か月後以内に作業計画を欧州委員会に提出しなければならない。CEN及びCENELECは、共同作業計画の修正を欧州委員会に通知しなければならない。」としている。

草案第２条４項はCEN及びCENELECに対して、「CEN及びCENELECは、欧州委員会に対し、全体的なプロジェクト計画へのアクセスを提供しなければならない。」としている。

４．５．標準化の実施期限

草案第３条は、①第１条の標準化要請の実施について、第２条の作業計画の実施の進捗状況を６ヶ月ごとに欧州委員会に報告すること、②第１条の標準化要請が採択されてから 10 ヶ月以内に、最初の共同半期報告書を欧州委員会に提出すること、③2025 年１月 31 日までに欧州委員会に共同最終報告書を提出すること、④附属書Ｉに定める期限及び要請の実施に関する重大な懸念事項があれば、すみやかに欧州委員会に報告すること、⑤第１項乃至第３項に基づく報告書は、EUの中小企業、市民社会組織の適切な関与及び関係者からの情報収集について、その計画及び実施内容に関する証跡を含めなければならないことについて定めている。

５．AI整合規格の要求事項

標準化要請案（草案）は、以下の 10 項目の要求事項を定めている。

【表４：10 項目の要求事項】

① AIシステムのリスク管理システム (Risk management system for AI systems)
② データとデータガバナンス (Data and data governance)
③ ログ機能による記録管理 (Record keeping through logging capabilities)
④ ユーザへの透明性と情報提供 (Transparency and information to the users)
⑤ 人間による監督 (Human oversight)
⑥ AIシステムの精度仕様 (Accuracy specifications for AI systems)
⑦ AIシステムの堅牢性に関する仕様 (Robustness specifications for AI systems)
⑧ AIシステムのサイバーセキュリティ仕様 (Cybersecurity specifications for AI systems)
⑨ 市販後モニタリング・プロセスを含む、AIシステム・プロバイダーのための品質マネジメントシステム (Quality management system for providers of AI systems, including post-market monitoring process)
⑩ AIシステムの適合性評価 (Conformity assessment for AI systems)

５．１．①AI システムのリスク管理システム

一つ目の要求事項は、「AI システムのリスク管理システム」である。

「本欧州規格又は欧州標準化規格類は、AI システムのリスク管理システムの仕様を定めなければならない。リスクマネジメントは、健康、安全又は基本的権利に関連するリスクを防止又は最小化することを目的とし、AI システムのライフサイクル全体を通じて実行される継続的な反復プロセスとして実施されるものでなければならない。」とし、「製品の安全構成要素である AI システムに該当する場合、AI システムに関連するリスクマネジメントシステムの要素が、製品全体のリスクマネジメントシステムに統合されるように、仕様を作成しなければならない。」としている。また、「仕様書は、関連事業者及び市場監視当局による利用の用に供することができるよう作成しなければならない。」としている。

５．２．②データ及びデータガバナンス

二つ目の要求事項は、「データ及びデータガバナンス」である。

欧州規格又は欧州標準化規格類を定めるにあたっては、以下の二つを含まなければならないとしている。

(a) AI システムのプロバイダが実施しなければならない適切なデータガバナンス及びデータ管理手順の仕様（特に重点を置く項目として、データの生成及び収集、データ準備作業、設計の選択、データの偏りやその他の関連する問題を検出して対処する手順）；

(b) AI システムの訓練、検証及びテストに使用されるデータセットの品質に関する仕様（代表性、関連性、完全性、正確性を含む）。

５．３．③ログ機能による記録管理

三つ目の要求事項は、「ログ機能による記録管理」である。

欧州規格において、「AI システムのイベントの自動ロギングに関する仕様を含むものと

する。これらの仕様により，システムのライフサイクル全体を通じて，システムのトレーサビリティ及び運用の監視が可能となり、提供者による AI システムの市販後の監視が容易になるものとする。」ことを要求している。

５．４．④ユーザへの透明性と情報提供

四つ目の要求事項は、「ユーザへの透明性と情報提供」である。

本欧州規格又は欧州標準化規格類において、次の事項に関する仕様を提供しなければならないとしている。

(a) ユーザがシステムの出力を理解し、それを適切に使用できるように、AI システムの運用の透明性を確保する設計及び開発ソリューション。

(b) AI システムに付随する取扱説明書であって、システムの能力及び限界に関する説明並びにメンテナンス及び保守に関する説明が含まれるもの。

(i) 専門家と一般のユーザ双方にとって適切で理解が容易な情報を識別し適切に区別する必要性；

(ii) (i)を損なうことなく、特定された情報が、利用者によるシステムの出力の解釈及び適切な利用を可能にするにあたって十分であることを保証する必要性。

５．５．⑤人間による監督

五つ目の要求事項は、「人間による監督」である。

本欧州規格又は欧州標準化規格類は、AI システムを人間が監視するための手段及び手順を定めなければならないとし、その対象を「(a) 技術的に実現可能である場合、AI システムの上市前又はサービスが提供される前に提供者によって特定され AI システムに組み込まれるもの」及び「(b) AI システムの上市又はサービス提供前に提供者が特定しユーザが実装するのに適切なもの」とし、「これらには、利用者が AI システムの運用に関連する事項を理解、監視、解釈、評価及び介入することを可能にする手段を含まなければならない。」としている。

また、本欧州規格又は欧州標準化規格類において、「特定の AI システムに特有の適切な監視措置も定義しなければならない。」とし、その対象として遠隔生体識別による本人確認を意図した AI システムに関しては、「人間の監視手段は、少なくとも２人の自然人によって個別に検証され確認されない限り、システムから得られる本人確認に基づいて利用者がいかなる行動又は意思決定も行わない可能性を予見しなければならない。」としている。

５．６．⑥AI システムの精度仕様

六つ目の要求事項は、「AI システムの精度仕様」である。

本欧州規格又は欧州標準化規格類は、「AI システムの適切な精度レベルを確保するための仕様を定め、提供者が関連する精度の指標とレベルを宣言できるようにするための仕様を定めること。」としている。また、「適切に定義されたレベルに対する精度を測定するための適切かつ関連するツール及び測定基準のセットも定義しなければならない。」としている。

５．７．⑦AI システムの堅牢性に関する仕様

七つ目の要求事項は、「AI システムの堅牢性に関する仕様」である。

本欧州規格又は欧州標準化規格類は、「関連するエラー、不具合及び不整合の原因並びに AI システムと環境との相互作用を考慮した AI システムの堅牢性に関する仕様を定めるものとする。」としている。

５．８．⑧AI システムのサイバーセキュリティ仕様

八つ目の要求事項は、「AI システムのサイバーセキュリティ仕様」である。

本欧州規格又は欧州標準化規格類は、「AI システムの脆弱性を悪意ある第三者が悪用することにより、AI システムの使用、動作又は性能を変更したり、セキュリティ特性を侵害する試みに対して、AI システムが弾力的であることを保証するために、適切な組織的及び技術的解決策を提供しなければならない。」としている。

組織的・技術的解決策として、「訓練データセット（データポイズニングなど）や訓練済みモデル（敵対的攻撃など）といった AI 固有の資産を操作しようとするサイバー攻撃や、AI システムのデジタル資産や基礎となる ICT インフラの脆弱性を悪用しようとするサイバー攻撃を防止・制御するための対策を必要に応じて定めなければならない。これらの技術的解決策は、関連する状況とリスクに適したものでなければならない。」としている。

さらに、当該仕様を定めるにあたっては、2022 年 9 月に欧州委員会により採択されたデジタル要素を有する製品に関するサイバーセキュリティの水平的要件に関する規則の提案の附属書 Ⅰ の第 1 節および第 2 節に記載されている、デジタル要素を有する製品に関する必須要件を十分に考慮しなければならないとしている。

５．９．⑨市販後モニタリング・プロセスを含む AI システム提供者のための品質マネジメントシステム

九つ目の要求事項は、「市販後モニタリング・プロセスを含む AI システム提供者のための品質マネジメントシステム」である。

本欧州規格又は欧州標準化規格類は、「AI の提供者が組織内において実施すべき品質マネジメントシステムの仕様を定めなければならない。」としている。

当該品質マネジメントシステムは、「AI システムが（2）、（3）、（4）、（5）、（6）、（7）及び（8）に記載された事項に継続的に適合することを確保しなければならないとし、中・小規模組織が品質マネジメントシステム対策を実施にあたって、適切な配慮がなされなければならない」としている。仕様書は、「AI システムに関連する品質マネジメントシステムが定める事項が、提供者の全体的なマネジメントシステムに統合され得るように作成されなければならない。」としている。

なお、品質マネジメントシステムの構築は、今後制定予定の ISO/IEC FDIS 42001 (Information technology - Artificial intelligence - Management system[46])に基づいて実施することになると想定されるが、欧州規格との整合性をどのように確保するのかは不明で

[46] ISO/IEC FDIS 42001 (Information technology - Artificial intelligence - Management system)<https://www.iso.org/standard/81230.html>.

ある。

また、品質マネジメントシステム認証は、ISO/IEC DIS 42006 (Information technology - Artificial intelligence - Requirements for bodies providing audit and certification of artificial intelligence management systems[47])に基づく認証が実施される予定である。

５．１０．⑩AI システムの適合性評価

最後の要求事項は、「AI システムの適合性評価」である。

本欧州規格又は欧州標準化規格類は、①AI システム及び AI 提供者の品質マネジメントシステムに関する適合性評価の手順及びプロセスを規定すること、②適合性評価業務を担当する者の能力を評価するための基準を提供することについて定めている。その際に、「適合性評価が提供者自身によって実施されるシナリオと、専門的な外部第三者機関の関与によって実施されるシナリオの両方を考慮しなければならない。」としている。

６．日本版 AI システム適合性評価制度構築に向けて必要な検討事項

６．１．DFFT（信頼性のある自由なデータ流通）との制度的な整合性の確保

AI システム適合性評価制度を構築した場合、AI システムの輸入、販売又はサービス提供を行う事業者に対する直接的な規制となるため、非関税障壁にあたるとの指摘を受ける可能性についても考慮が必要である。そのため、DFFT（信頼性のある自由なデータ流通）との制度的な整合性の確保が重要である。

DFFT（データ・フリー・フロー・ウィズ・トラスト）は、IT 総合戦略本部「デジタル時代の新たなＩＴ政策大綱」（2019 年６月７日決定）において示されたものである。IT 政策大綱は、「プライバシーやセキュリティ・知的財産権に関する信頼を確保しながら、ビジネスや社会課題の解決に有益なデータが国境を意識することなく自由に行き来する、国際的に自由なデータ流通の促進を目指す」というコンセプトを掲げ、「DFFT（信頼性のある自由なデータ流通）のコンセプトに基づく「国際データ流通網」を広げていくことを目的として、より多くの国との間で、デジタル貿易ルールの形成等を促進することが求められる」とし、当該目的を達するため、「世界最先端デジタル国家創造宣言・官民データ活用推進基本計画」（2019 年６月 14 日閣議決定）が策定された。

G20 茨城つくば貿易・デジタル経済大臣会合（令和元年６月８日及び９日）において、DFFT は G20 全体で合意され、信頼につながる各国の法的枠組みは相互に接続可能なものであるべきことが確認された。その後、2019 年の G20 大阪首脳宣言、2021 年の「DFFT への協力に向けた G7 ロードマップ」、2022 年の「DFFT 促進のための G7 アクションプラン」と着実に取り組みがなされている。

情報の自由な流通と保護を基調とするデータ保護制度の整備は、1980 年に OECD プライバシーガイドラインが制定され、OECD 加盟国相互における個人データの自由な流通と保護を目的とする基本理念が共有され培われてきたものである。この理念は、我が国においては、個人情報保護法第１条が、「個人情報の適正かつ効果的な活用が新たな産業の創出並

[47] ISO/IEC DIS 42006(Information technology — Artificial intelligence — Requirements for bodies providing audit and certification of artificial intelligence management systems)<https://www.iso.org/standard/44546.html>.

びに活力ある経済社会及び豊かな国民生活の実現に資するものであることその他の個人情報の有用性に配慮しつつ、個人の権利利益を保護することを目的とする。」と定め、個人情報の適正な「利用」と「保護」のバランスのもとで個人情報を取り扱う事業者の遵守すべき義務を定めている。

DFFT に基づく「国際データ流通網」の拡大を目的として、多国間のデジタル貿易ルールの形成の促進を目指す施策としては、「デジタル貿易に関する日本国とアメリカ合衆国との間の協定（略称：日米デジタル貿易協定）」（令和２年１月１日）」、「日英包括的経済連携協定（EPA）」(令和３年３月 26 日)」、「TPP11（環太平洋パートナーシップに関する包括的及び先進的な協定）」等の二国間・複数国間の貿易協定の電子商取引に関する取り決め部分に、DFFT に関連する規律が含まれている。

６．２．DFFT 及びデータガバナンスと AI 規制の両立

G7 広島サミットの「閣僚宣言：G7 デジタル・技術大臣会合（2023 年４月 30 日）」において、DFFT 及びデータガバナンスは「越境データ流通及び信頼性のあるデータの自由な流通の促進」、AI 原則や信頼できる AI に関する施策は「責任ある AI と AI ガバナンスの推進」と両者の項目は分かれている。

DFFT の枠組みは、AI システム適合性評価制度にも適用できる。その理由は、DFFT を実現するための国際的なトラスト基盤の構築は、「標準化」と「認定・認証」により達成されるものであり、本稿が提案する適合性評価制度と目標も構成も一致する。適合性評価制度におけるデータ・ガバナンスについては DFFT の枠組みに基づく管理方策を提案することが可能と考えられる。

DFFT と AI ガバナンスに向けた取り組みは統合されておらず別個の施策として並行している。この併存を解消し一体化させることで、日本の強みである DFFT の一層の促進を図りつつ、AI ガバナンスに向けた取り組みについて DFFT を礎として展開することができる。

６．３．地政学的観点からの AI 規制

DFFT は、データ管理を一定の域内に限定する「データローカライゼーション」と対照的な概念として提唱されたものである。各国の AI 戦略は、米国のプラットフォーム事業者による AI の利活用、EU における AI 規制に向けた取り組み、中国などのアジア諸国との関係における経済安全保障の観点から様々な検討がなされつつある。新たな先端技術である AI などの機微技術管理は、経済安全保障の枠組みにおいて保護する方針も示されている。

我が国においても、国際情勢の複雑化、社会経済構造の変化等により、安全保障の裾野が経済分野に急速に拡大していることから、国家・国民の安全を経済面から確保するための取組を強化・推進するため、「経済施策を一体的に講ずることによる安全保障の確保の推進に関する法律」（経済安全保障推進法）（令和４年法律第 43 号）が制定されている。

経済安全保障推進法では、安全保障の確保に関する経済施策として、①重要物資の安定的な供給の確保、②基幹インフラ役務の安定的な提供の確保、③先端的な重要技術の開発支援、④特許出願の非公開、の４つの制度を創設している。支援対象となる先端的な重要技術は、宇宙・海洋・量子・AI・バイオ等の分野における先端的な重要技術の研究開発と成果が

対象となっている。

改正前の個人情報保護法第24条に係る委員会規則の方向性に冠する検討段階の資料[48]では外国指定の要件を提示しており、５つの要件のうち、④相互の理解、連携及び協力が可能であること、⑤個人情報の保護を図りつつ相互の円滑な移転を図る枠組みの構築が可能であること、を定めている。これらの二つの要件は、相互理解に基づくデータ移転が可能なDFFTの枠組み設定の理念と一致するものである。

国際標準や適合性評価制度は経済安全保障の観点からの検討も必要であるが、国際標準はDFFTの理念とは異なり参画要件はあくまで当該標準への「準拠」のみであり、相互理解などの制限を設けることができない。ゆえに、AI関係国際標準の策定は、経済安全保障を確保する観点からは参入を拒むなどの対策を講じることができない点が課題となることに留意が必要である。

７．適合性評価制度の構成要素

７．１．EUのAI法案が定める適合性評価制度の構成要素

EUのAI法案が定める適合性評価制度は、①認証機関（30-39条）、②適合性評価手続（40-51条）、③高リスクAIシステムのマネジメントシステム要求事項（８-15条)、④CEマークの表示から成る。適合性評価の方法は、整合規格が存在する場合の要件適合（40条）と整合規格不存在の場合の共通仕様策定（41条）から成る。

この仕組みを参考に、我が国においてAIシステム適合性評価制度を構築するにあたって必要な構成要素は、これまでマネジメントシステム規格を整備するにあたって検討が行われてきた事項と基本的に同じであることがわかる。

７．２．日本版AIシステム適合性評価制度構築にあたって必要な構成要素

日本版AIシステム適合性評価制度を検討するにあたっては、まずは根拠法において対象となるAIシステムの定義が必要である。EUの場合は高リスクAIを対象としていることから我が国においてもその対象範囲を定める必要がある。EUのAI法案が定める適合性評価制度に照らし、以下、日本国内において実施する場合の各構成要素を検討する。

①適合性評価の実施機関及び②評価手続については、JIS Q 17011：2018（ISO/IEC 17011：2017)「適合性評価 - 適合性評価機関の認定を行う機関に対する要求事項」に基づいて設置することができる。ただし、JIS Q 17011の対象となる適合性評価機関は、試験所・校正機関・製品認証機関であり、JIS Q 17025「試験所及び校正機関の能力に関する一般要求事項」、JIS Q 17034「標準物質生産者の能力に関する一般要求事項」、JIS Q 17065「適合性評価―製品、サービス及びプロセスの認証を行う機関に対する一般要求事項」からなる。そのため、AIシステム適合性評価制度の実施に用いるためにはJIS Q 17065の改定かAIの適合性評価制度に特化したマネジメントシステム規格の新設が必要となる。

③のマネジメントシステム規格に基づく実施内容の審査、評価及び認証を行う手続については、JIS Q 17021-1：2015（ISO/IEC 17021-1：2015）「適合性評価-マネジメントシス

[48] 個人情報保護法第 24 条に係る委員会規則の方向性について
<https://www.ppc.go.jp/files/pdf/290616_siryou1.pdf>。

テムの審査及び認証を行う機関に対する要求事項-第 1 部：要求事項」に基づいて実施することができる。なお、日本国内には様々な認証機関が存在するが、製品安全規制とデータ管理のいずれの枠組みから実施機関の設置を検討するのかにより、指定される認証機関及びその評価手続は異なることになる。また、民間認証によって事実上の強制規格としての AI システム適合性評価制度を実現することは困難であると考えられるため、認証機関の設置については、後述の AI 規正委員会の所掌とすべきである。

②及び③については、現行の JIS 規格の活用では限界があるため、「AI の適合性評価制度に特化したマネジメントシステム認証」を実施するため、産業標準化法に基づく新たな「AI マネジメントシステム規格（仮称）」の制定を検討する案を提案したい。その理由は単に現行 JIS の要求事項では不十分ということだけではなく次の二つの理由がある。

一つ目は、マネジメントシステム規格の共通基本構造を定めている ISO/IEC 専門業務用指針 補足指針 2023 年（第 3 版）[49]の附属書 SL[50]と、EU の AI 法案の高リスク AI に係るマネジメントシステムの内容を比較すると附属書 SL とは不整合である点があげられる。つまり、我が国において AI システム適合性評価制度におけるマネジメントシステム規格を策定する際に附属書 SL との整合性を確保した要求事項で構成する規格を策定することで、国際標準としての AI システム適合性評価制度を提案することが可能になる。

ただし、AI に関する国際標準の検討は、2017 年に ISO/IEC JTC1/SC42 が設置され、①AI ガバナンス、②基礎的標準、③データ、④信頼性、⑤ユースケースと応用、⑥計算アプローチと計算的特徴の 6 つの WG における検討が進んでいる。前述の通り、ISO/IEC FDIS 42001 (Information technology - Artificial intelligence - Management system[51])が今後策定されると当該規格は JIS 化される予定であることから、国内規格の検討よりも国際標準の策定のほうが先行するため、マネジメントシステム規格とともに認証規格についても、ISO が策定する AI 国際標準規格を JIS 化する方向が現実的であることは否めない。

二つ目は、これまで検討がなされてきたガイドラインや指針等に定められている「原則」をはじめ、AI の研究開発から利用において留意すべき実体的な利益の保護のための規定は、マネジメントシステム規格の「附属書」として編入することが可能である。AI システム適合性評価制度における手続的な義務とともに、精緻な議論がなされてきた個人の権利利益

---

49 ISO/IEC Directives, Part 1 Procedures for the technical work — Consolidated ISO Supplement — Procedures specific to ISO Edition, 2023 (ISO/IEC 専門業務用指針，第 1 部 専門業務の手順 –統合版 ISO 補足指針 – ISO 専用手順 2023 年（第 3 版）)<https://webdesk.jsa.or.jp/pdf/dev/md_6032.pdf>。

50 マネジメントシステム規格の整合化動向 <https://webdesk.jsa.or.jp/common/W10K0500/index/dev/iso_mngment03/>。附属書 SL の詳細は、前掲資料 136 頁内に記されている URL から、そのガイダンス資料である Annex SL Appendix 2(normative)Harmonized structure for MSS with guidance for use にアクセスし参照されたい。

51 ISO/IEC FDIS 42001 (Information technology - Artificial intelligence - Management system)<https://www.iso.org/standard/81230.html>.

保護や原則の遵守を組み込んだ基本理念に基づく制度の構築を目指すべきであろう。

④AI システム適合性評価制度に基づく認定を受けたことを示すマーク制度又はシールプログラムは、電気用品安全法に基づく PSE マークの活用が考えられるが、当該マークの対象は「電気製品」である。EU の AI 法案の CE マークと対比すると、EU の機械指令（機械規則提案）に対応する部分については整合するものの、「AI サービス」への適用については課題が残る。さらに、AI システムをネットワークに接続する場合は、端末機器の技術基準適合認定と無線通信については技術基準適合証明を取得する必要がある。したがって、(a)電気用品安全法の改正により PSE マークを AI システム適合証明に活用する方法を検討し、既存の「技術基準適合認定・証明 [52]」は現行法の枠組みで対応するのか、(b)日本版の AI 新法を制定し規制対象の AI の定義を定めるとともに、整合規格への適合から表示（マーク制度）に至る義務を法定し、整備法において適合認定・証明手続を一括して定めハネ改正を行うのか、(c)産業標準化法に基づく「AI マネジメントシステム規格（仮称）」の認証制度に基づく新たなマーク制度を創設するのか、様々な案が考えられる。

なお、産業標準化法第 30 条第 1 項その他の関係規定に基づく「JIS マーク [53]」のようなマークと異なるところは、対象となる AI システムの製造、販売やサービス提供にあたって認定を受けることが義務付けられる「強制規格」を念頭に置いている点である。

７．３．「AI マネジメントシステム」と「AI マネジメントシステム規格」

「AI マネジメントシステム」という名称を用いているものとして、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. 1.0[54]」が示す「健全な事業活動倫理を尊重するためのマネジメントシステム」がある。ガイドラインでは、「AI の社会実装の促進に必要な AI 原則の実践を支援すべく、AI 事業者が実施すべき行動目標を提示するとともに、それぞれの行動目標に対応する仮想的な実践例や AI ガバナンス・ゴールとの乖離を評価するための実務的な対応例も例示」し、「AI システムの開発・運用等に関わる事業者の取引等で広く参照されることや、AI 原則の実践に関するステークホルダーの共通認識の形成を通じて、各社の自主的な取り組みを後押しすることが期待される」ものであるとしている [55]。

---

[52] AI システムについて通信端末機器を電気通信事業者のネットワーク（電気通信回線設備）に接続し使用する場合は、電気通信事業法 52 条第 1 項に基づく「端末機器の技術基準適合認定等に関する規則」（平成 16 年総務省令第 15 号） による「技術基準適合認定」の対象となる。上市規制との関係では、輸出入に係る端末機器の技術基準適合認定は、欧州共同体（EC）（平成 14 年 1 月発効）、シンガポール（平成 14 年 11 月発効）及び米国（平成 20 年 1 月発効）との間で MRA（相互承認協定）がある。無線通信を行う特定無線設備については、電波法第 38 条の 2 に基づく「特定無線設備の技術基準適合証明等に関する規則（昭和 56 年郵政省令第 37 号)」による「技術基準適合証明」を取得する必要がある。

[53] JIS マーク<https://www.jisc.go.jp/newjis/newjismknews.html>.

[54] 経済産業省：AI 原則の実践の在り方に関する検討会：AI ガバナンス・ガイドライン WG「AI 原則実践のためのガバナンス・ガイドライン ver. 1.0」（令和３年７月９日）。

[55] 前掲３頁。

これまでも、AI を企業活動において利用するにあたり、そのガバナンスを達成するために企業内部における原則や指針策定の取り組みがなされてきたが、それらは、企業活動におけるコーポレートガバナンスや内部統制の一環としてのAIガバナンス体制の整備を目的とするものである。そのため、そのような原則や指針等の活用はもっぱら企業の自主的な取り組みに委ねられている。このようなガイドラインを遵守する企業は、AI ガバナンスへの取り組みの必要性を認識するほど意識が高く、具体的に取り組みを行うことが企業活動において有益であると判断し、実際に取り組みを実施するだけの余裕と能力を有する企業に限られる。実際に、その取り組みを行っている企業は、かなり高度なAIガバナンスに向けた取り組みを実施している企業であるといえる。このようなガイドラインを遵守することは、企業活動におけるAIガバナンス達成のための高尚な取り組みにおいて指針となるものではあるものの、そのような意識が高くない企業においては単なる参考資料としての位置付けにとどまり、具体的な取り組みに着手することは難しいと考えられる。

以上から、現時点で我が国において示されているマネジメントシステムの構築に向けた取り組みは、EU の AI 法案が定める高リスク AI について CE マークを付すための整合規格の実施のためのマネジメントシステムの構築とは目的も趣旨も異なるものである。

ゆえに、本稿の提案に係るマネジメントシステムとは、AI システム管理のための「マネジメントシステム規格」のことをいう。

８．AI 規正委員会（仮称）の設置

本稿が提案する日本版 AI システム適合性評価制度を機能させ、民間部門とともに公的部門による AI の利用についても監督を行うための組織として、国家行政組織法第三条に基づくいわゆる三条機関（委員会）として、「AI 規正委員会（仮称）」の設置を提案したい。

EU の AI 法案が定める欧州人工知能委員会(EAIB)に対応する機関の設置が求められるところであるが、名称については、AI 規制(regulation)のための監督(supervisory)及び管理(management）を行う第三者機関に求められているのは AI ガバナンスであるから、AI 統治(governance）委員会が適切であると考えられるが、その名称はさておき第三者機関設置の必要性が提案の本旨であることから、単に「AI 規正委員会」とした。

なお、サイバネティック・アバターその他の新興技術の適合性評価制度への拡充という観点からすると、「AI」の名称を冠し AI 技術に特化した委員会ではない体制も視野に入れる必要がある。その場合は、例えば、米国が進めている８分野（通信・ネットワーク技術、半導体、AI・機械学習、バイオテクノロジー、測位・航法・タイミング（PNT）サービス、デジタルアイデンティティ・インフラ及び分散台帳技術、クリーンエネルギー発電と蓄電、量子情報技術）の重要・新興技術（Critical and Emerging Technologies(CET)）に係る戦略と歩調を合わせ、「重要・新興技術委員会」とすることも考えられる。

AI システム適合性評価制度を機能させるための認証制度については、民間組織による認証制度は国際的な相互認証を実現する上で支障となることがある。そのため、AI 規正委員会に認証部門を設け、事実上の強制規格として実施することが望ましい。なお、同様の趣旨について、個人情報保護法の３年毎見直しにおいて、民間認証である「プライバシーマーク

制度」を個人情報保護委員会に認証部門を設置して実施すべきであるという見解[56]を述べた。その趣旨は、国際的な個人情報保護の枠組みにおいてJIS Q 15001を国際標準として機能させ、さらにEUの一般データ保護規則(GDPR)42条のシールプログラムとの相互認証を実現することにあった。GDPRに基づくシールプログラムとの相互認証については、その協議に向けた兆候はあったものの、日本国内におけるシールプログラム（マーク制度）の実施が民間組織によるものであることから、現在に至るまで実現には至っていない。

ゆえに、事業者側が実施するマネジメントシステム構築・運用指針である「AIシステムマネジメント規格」を産業標準化法に基づいて制定する場合、当該規格への適合性の審査基準は、行政手続法５条の審査基準及び同12条の処分基準に基づいて、AI規正委員会が整合規格を策定し要求事項を定めるべきである。

AIの研究開発の促進など、EUのAI法案53条が定める「サンドボックス」の活用など、我が国では国家戦略特区及び構造改革特区を活用したAI関連イノベーションの促進も重要な施策であることから、AI政策全般の実施に必要な施策推進が重要な所掌事務であるという観点も重視すべきである。

なお、EUのAI法案56条は、①国内監督機関と欧州委員会の実効的な協力への貢献、②これらの機関の指針等の調整、③これらの機関の補助をEAIBの役割と規定し、委員会の役割を欧州委員会に対する助言及び補助としている[57]。同59条に基づいて、「職務権限のある国内機関の指定」がなされるが、2023年８月22日に最初の機関としてスペインがNational Artificial Intelligence Supervisory Agency[58]を設置している。

９．おわりに

非拘束的なガイドラインを軸とする自主的な規律に基づく我が国における取り組みに対し、(a)法規制回避論からの脱却、(b)AI規制の最適化、(c)日本「発or初」の新たなAI規制政策立案の観点から、①AI規制の導入、②自主性の尊重、③事実上の強制規格の導入、④AIマネジメントシステム規格の策定、⑤AI規制に係る根拠法の整備、⑥新たな監督機関の設置、⑦整合規格に準拠する技術基準適合義務を課し適合証明の表示を義務付ける「日本版AIシステム適合性評価制度」から構成される新たなAI規正を提案した。

非拘束的なガイドラインに基づく自主的な規律は、それに自主的に賛同する組織に対してはその取り組みを推進する上での指針となるため有効である。しかし、規制の根拠となる法令が存在しない状況では、準則や法解釈の指針としてのガイドラインではなくガイダンス的な機能を発揮しているにすぎない。

一方で、今後、日本企業だけでなく各国のAIシステム開発・販売・提供を行う組織は、EUの適合性評価制度に強制的に従わざるを得なくなる。既に我が国においても、EUの適

---

[56] 第105回 個人情報保護委員会（令和元年５月17日（金））資料４　新保史生提出資料<https://www.ppc.go.jp/aboutus/minutes/2019/20190517/>。

[57] 夏井・前掲注34、491-493頁。

[58] Spain: Council of Ministers approves Agency for the Supervision of AI<https://www.dataguidance.com/news/spain-council-ministers-approves-agency-supervision-ai>(2023.8.22)

合性評価制度への対応を念頭に置いた事業展開を目指す事業者も現れている。日本企業としても、AI システムを開発し販売を企図する場合に EU 市場を放棄するわけにはいかず、事実上の世界標準として機能する EU の適合性評価制度に準拠した対応を迫られる。その際に、法令遵守意識が高い日本企業は、忠実にその制度に準拠するための取り組みを進めるであろう。強制的に遵守せざるを得ない規格や制度への準拠については従順に対応し単にそれらの規制に盲従するしかない状況を、AI 規制への対応においては見直さなければならない。法規制を回避し自主的な規律の推進を掲げた政策を継続し、国際動向を注視していながら気が付くと国際的には法規制に舵を切った政策が主流になった時、AI 政策分野での日本の凋落を見守るしかない状況に陥るおそれがある。

本稿で示した AI 規制に向けた新たな構想が、日本の AI 規制政策の隘路を認識するきっかけとなることを願いたい。

本研究は、JST ムーンショット型研究開発事業 [59]、JPMJMS2215 の支援を受けたものである。

[59] 本稿を執筆するにあたっては、ムーンショット研究開発プロジェクト目標 1 「アバターを安全かつ信頼して利用できる社会の実現」の研究参加者であり、ISO/IEC JTC1/SC42 (AI)専門委員会幹事で慶應義塾大学大学院政策・メディア研究科特任教授の江川尚志氏から欧州の標準化動向について情報提供を得た。