

AI 規正論 2.0 —適合性評価制度の展開と統合認証の実現に向けて

新保 史生¹ (慶應義塾大学)

要 旨

本稿は、2023年に提唱した「AI 規正論」を発展させ、AI システム適合性評価制度の展開と統合認証の実現に向けた考察を行うものである。EU では AI 法に基づく整合規格策定の遅延やオムニバス提案による規制緩和の動きが見られ、ブリュッセル効果の変容も顕著である。一方、日本では AI 制度標準化の取り組みとして JIS Q 42001 が制定され、マネジメントシステム規格に基づく AI システムの組織的管理に関する枠組みが整備されている。国際標準化においては、ISO/IEC JTC 1/SC 42 での日本の貢献が大きく、複数のワーキンググループで主導的役割を担っている。AI 規正論 2.0 の目標は、製品認証とマネジメントシステム認証を統合した「統合認証」の実現である。サイバネティック・アバターの実現研究を通じた実用性の検証とともに、統合認証の枠組みを提示し信頼できる AI 社会の実現に向けて、AI 統治のベストプラクティスを世界に発信していくことにある。

キーワード：AI 統治、統合認証、適合性評価、マネジメントシステム、サイバネティック・アバター

1. はじめに

1. 1. AI 規正論 1.0 から 2.0 へ

「AI 規正論²⁾」は、2023年に本誌『情報通信政策研究』7巻1号³⁾において提唱した構想である(以下、「AI 規正論(1.0)」という)。その端緒は、2021年にEUの「AI 整合規則提案(harmonised rules on artificial intelligence)(AI 法案)⁴⁾」が公表されたことにある。

¹ 慶應義塾大学総合政策学部教授

² 「規制」は、法令等により行為を制限・禁止することを主眼とする用語である。一方、AI 規正論 1.0 から用いている「規正」は、「規(のり)」すなわち基準・準則を定め「正(ただ)す」ことを意味し、単なる制限ではなく適切な基準への適合を通じて秩序を正すという積極的な意味を含む。本稿が提案する AI 統治の枠組みは、AI のリスクを殊更に強調して禁止・制限するための検討を行うのではなく、整合規格への適合、マネジメントシステムの構築、第三者認証といった仕組みを通じて、AI システムの信頼性と安全性を担保し、社会における適正な利用を実現することにある。この趣旨を表すため、「規制」ではなく「規正」の語を用いている。

³ 新保史生「AI 規正論」『情報通信政策研究』7巻1号 69-100頁(2023年)。

⁴ 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.

る。

EU では域内の AI 規制の整合化を図るため「AI 整合規則 (AI 法)⁵」を制定し、高リスクに分類される AI システムに製品安全規制同様の義務を課して CE マーキングの対象とするため、適合性評価及び第三者認証制度を整備する制度の構築が進められてきた。

しかしながら、このような製品安全規制の枠組みは、日本が Made in Japan の信頼性を培う中で長年にわたり実践してきた仕組みに他ならない。我が国の産業競争力は、電気・電子製品をはじめとする様々な分野において、ISO 規格・IEC 規格等の国際標準及びこれらに整合する JIS 規格に準拠し、安全かつ安心して利用できる信頼性ある製品の開発・販売によって実現されてきた。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 (QMS)、環境 (EMS)、労働安全衛生 (OHSMS)、事業継続 (BCMS) などの認証制度、JIS Q 15001 を用いたプライバシーマーク制度、JIS Q 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。

すなわち、EU が AI 法により構築しようとしている製品安全規制、適合性評価、第三者認証の枠組みにおいて、日本は既に豊富な実績と経験を有しており、むしろ優位な立場にある。AI 規正論 1.0 は、この優位性を認識した上で EU の AI 整合規則提案 (AI 法案) を分析⁶ し、日本が AI 統治の分野において主導的役割を果たし得ることを提示するとともに、日本版の AI 整合規格を策定し適合性評価制度を整備・展開すべきであると提案したものである。

1. 2. AI 規正論 2.0 の構成と目的

AI 規正論 2.0 は、①EU の AI 制度標準化、②日本の AI 制度標準化、③国際標準の検討状況、④統合認証の必要性の 4 つの観点から、AI ガバナンス (以下、「AI 統治」という。) の動向を把握し、AI システム適合性評価制度の展開と統合認証の実現に向けた考察を行うものである。

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、安心して使うことができることである。「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

⁵ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), OJ L 2024/1689. EU の AI 法の正確な理解のためには、夏井高人教授の翻訳と解説を参照されたい。夏井高人 (訳) 「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)。夏井高人 (訳) 「規則(EU) 2024/1689 (人工知能法) [参考訳]」法と情報雑誌 9 巻 3 号 1-130 頁 [前文]、同 9 巻 4 号 1-143 頁 [条文第 1 条~第 63 条]、同 9 巻 5 号 1-88 頁 [条文第 64 条~第 113 条]、同 9 巻 5 号 89-145 頁 [別紙] (法と情報研究会、2024 年)。

⁶ 新保史生「AI 規制の国際動向」都市問題 115 巻 2 号 18-25 頁 (2024 年)。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 61-80 頁 (日本経済研究センター、2023 年)。

現するものであり、本稿の提案意図は、AI 時代に向けて必要な AI の管理及び規制の方策として新たな制度設計を提示することにある。

2. EU の AI 制度標準化

2. 1. 適用開始スケジュール

EU の AI 法は、2024 年 5 月 21 日の EU 理事会承認後、表 1 の当初のスケジュールで適用が開始される予定であった。

本稿執筆時点において、「一般目的 AI⁷ (GPAI: General Purpose AI⁸)」に関する規定は既に適用が開始されているが、それ以降の適用開始については、①標準化要請に基づく整合規格 (Harmonised Standards) の策定の遅延、②適合性評価を実施する指定組織 (Notified Body) による認証体制の未整備、③欧州理事会が「設計段階からの簡素化 (simplicity by design)」に基づく「EU 規制の簡素化⁹」を提案し、2026 年 8 月以降に予定されていた適用開始スケジュールは延期が予定されている。

【表 1: AI 法の発効と段階的な適用開始】

適用開始日	項目
2024年8月1日	発効 (EU官報公示の20日後)
2025年2月2日	第I章 (総則条項) 及び第II章 (禁止されるAI実務) の適用開始

⁷ 一般目的 AI (GPAI) の定義は、AI 法第 3 条 63 号の「一般目的 AI モデル」(大量のデータを用いた大規模な自己教師あり学習等により訓練され、広範囲の異なるタスクを適切に実行する能力を有し、様々なシステムやアプリケーションに統合可能な AI モデル) と、同 66 号の「一般目的 AI システム」(一般目的 AI モデルを基礎とする直接利用と他の AI システム中への組込む AI システム) からなる。

⁸ GPAI: General Purpose AI の訳について、「General」は、GDPR の「General」と同義である。よって、法令用語の訳語の一貫性を確保する観点からは、GDPR は「一般」と訳出してきたことから適切な訳語は「一般目的 AI」であり、「汎用目的 AI」という訳語ではこれまでの訳語との一貫性を欠く。なお、汎用 AI (AGI: Artificial General Intelligence) ではないので、GPAI を「汎用 AI」と訳すことは誤りである。なお、AI 法第 56 条に基づき、GPAI 提供者が AI 法上の義務遵守を示すための具体的な実務指針が「実務準則 (Code of Practice)」であり、同法第 95 条に基づく、より広範な任意の行動規範である「行動準則 (Code of Conduct)」とは異なるものである。また、GPAI 実務準則は任意の取り組みであり、AI 法の遵守状況を確認する一つ的手段ではあるが、AI 法準拠に向けた義務ではない。European Commission, *General-purpose AI models in the AI Act: Questions and Answers*, last modified October 2025, <<https://digital-strategy.ec.europa.eu/en/faqs/general-purpose-ai-models-ai-act-questions-answers>>.

⁹ European Council, *Simplification of EU rules* <<https://www.consilium.europa.eu/en/policies/simplification>>, (2026 年 2 月 1 日確認)。

2025年8月2日	以下の章の適用開始 第III章第4節（通知元当局及び指定組織） 第V章（一般目的AIモデル） 第VII章（統治） 第XII章（罰則）※第101条を除く第78条（機密性）
2026年8月2日 （2027年12月2日への最大16ヶ月の延期予定）	規則の全面適用 第III章第1節（高リスクAIシステムの分類） 第III章第2節（高リスクAIシステムに関する要件） 第III章第3節（高リスクAIシステムの提供者と設置者及びその他関係者の義務） 第IV章（一定のAIシステムの提供者及び設置者の透明性義務）（第50条第2項のAI生成コンテンツの機械可読検出について、2025年8月2日より前に市場に置かれた生成AIシステムには2027年2月2日までの猶予期間が設定される予定） 第VIII章（高リスクAIシステムのEUデータベース） 第IX章（市場化後監視、情報共有及び市場監視） 第X章（行動準則及び指針） 第XI章（権限の委任及び委員会の手続） 第XIII章（最終規定）
2027年8月2日 （2028年8月2日への最大12ヶ月の延期予定）	第6条第1項及びそれに対応する義務（高リスク製品の安全性コンポーネントとしてのAIシステムに関する規定）の適用開始
2025年8月2日まで	構成国による単一連絡部局の公表（第70条第2項）
2026年8月2日まで	構成国による少なくとも1つの国内レベルのAI法制サンドボックスの開設（第57条第1項）
2027年8月2日まで	2025年8月2日より前に市場に置かれた一般目的AIモデルの提供者による規則遵守のための措置実施（第111条第3項）
2030年8月2日まで	行政機関によって利用されることが意図された高リスクAIシステムの提供者及び設置者による規則遵守のための措置実施（第111条第2項）
2030年12月31日まで	別紙Xの大規模ITシステムのコンポーネントであるAIシステムの規則遵守（第111条第1項）

2. 2. 一般目的 AI モデルに関する規律の性格

適用開始対象の規定のうち、第V章が定める一般目的AIモデル（以下「GPAIモデル」という。）に関する規律の性格には注意を要する。高リスクAIシステムの提供者には、適

合性評価（第 III 章第 5 節）や CE マーキング（第 48 条）に関する義務が課される。これに対し、GPAI 提供者に適用される第 V 章の規律には強制的な適合性評価の仕組みが存在しない。AI 法第 53 条及び 55 条は条文上「義務（Obligations）」と題しているものの、その遵守の実証手段は GPAI 提供者の判断に委ねられている。具体的には、同法第 56 条に基づき策定された「一般目的 AI 実務準則」（GPAI Code of Practice）への任意の署名、又は提供者が独自に選択する代替的手段のいずれかにより適合を示せば足りる。欧州委員会は当該実務準則を「適切な任意のツール（adequate voluntary tool）」と位置づけている。すなわち、GPAI 規律は、第三者による強制的な適合性評価を基盤とする高リスク AI システムの規制枠組みとは本質的に異なり、提供者の自主的な対応に依拠した枠組みである。

もっとも、AI 法は、GPAI モデルのうち「システムミック・リスク」を有するものについては強化された規律を設けている。同法第 3 条 65 号が定めるシステムミック・リスクに関する規定を要約すると、GPAI モデルの高影響能力（high-impact capabilities）に起因し、EU 市場や公衆衛生・安全・基本的権利等に広範な負の影響を及ぼし得るリスクであって、AI のバリューチェーン全体に伝播し得るものをいう。化学兵器・生物兵器の開発障壁の低下や、自律的に動作する GPAI モデルの制御喪失等がその具体例として AI 法の前文に例示されている。

AI 法第 51 条 1 項は、GPAI モデルがシステムミック・リスクを有するモデルとして分類される条件を定め、同条 2 項は、その推定基準として、訓練に使用された計算量の累積値が浮動小数点演算（FLOPs: Floating Point Operations）¹⁰で 10^{25} を超える場合を規定する。この閾値を超えるモデルの提供者には、モデル評価、敵対的テスト、重大インシデントの報告、サイバーセキュリティの確保等が同法第 55 条で求められるが、その遵守方法もまた、実務準則への任意の署名又は代替手段の提示に委ねられている。ただし、欧州委員会はこの閾値を委任法令により改定する権限を有しており（第 51 条 3 項）、対象範囲は技術的進展に応じて変動し得る。

このシステムミック・リスクに関する解釈には、少なくとも複数の解説¹¹で混同が確認できる。AI 法第 51 条 2 項が定める「 10^{25} FLOPs(10 の 25 乗 FLOPs 訓練計算量閾値)」の解釈が一例であり、その誤解や FLOPs と FLOPS の混同は規制対象の範囲を誤認させる深刻な問題となり得る。

EU の AI 法第 51 条 2 項が定めているのは FLOPs（Floating Point Operations）であ

¹⁰ AI 法第 51 条 2 項は、訓練に用いられた計算量の累積（cumulative amount of computation used for its training）を、floating point operations で測定した値が 10^{25} を超える場合に、高い影響力の能力（high impact capabilities）を推定するとしている。

¹¹ システムミック・リスクに関する解釈を混同している一例として、中国のオープンソース AI モデル「DeepSeek」の登場により、欧州委員会が AI 法における一般目的 AI モデルの分類基準、特に FLOPS 閾値の見直しを検討しているとし、AI 法第 51 条 3 項に基づき、欧州委員会は「システムミック・リスクを伴う」GPAI モデルの分類と閾値を最新の技術・業界慣行を反映するよう更新する法的義務を負っていることを解説している。Out-Law (Pinsent Masons), EU AI Act rules on GPAI models under DeepSeek review (17 Nov. 2025), <<https://www.pinsentmasons.com/out-law/analysis/eu-ai-act-gpai-deepseek-review>>.

り、これは AI モデルの訓練に使用された計算量の総和を指す。一方、FLOPS は Floating Point Operations Per Second の略であり、1 秒あたりの演算速度を意味する。EU の AI 法が基準としているのは演算の「累積量」であり「速度」ではない。

本稿執筆時点で 10^{25} FLOPs 訓練計算量¹²に達している日本の国産 AI モデルは存在しないが、この基準を誤解したまま「当面規制対象にならない」と判断していた企業が、今後の技術開発の進展によりシステミック・リスクを引き起こした際に規制対象になっていることが判明した場合、それに伴う制裁金リスクは計り知れない。GPAI 規律の遵守方法が提供者の自主的対応に委ねられているとはいえ、規律そのものが存在しないわけではない。GPAI 提供者に対して全世界年間売上高の 3%又は 1,500 万ユーロのいずれか高い方を上限とする制裁金を定めている AI 法第 101 条の規定は、その適用が 2026 年 8 月 2 日から 2027 年 12 月 2 日に延期になる見込みであるものの、遵守方法の任意性は制裁の不在を意味しない。

2. 3. AI 法の目的と適合性評価制度

EU の AI 法の目的の中核にある安全規制の枠組みは、AI システム適合性評価制度の導入を目指す新たな「整合規則 (harmonised rules)」の制定である。AI 法は単に AI システムの規制を目的とするだけでなく、既存の製品安全制度と AI 技術を統合し、包括的な製品安全適合性評価制度を構築しようとする野心的な規制である。

適合性評価の仕組みにおいては、Notified Body (指定組織) の設置がある。これは第三者が品質を判断する仕組みであり、AI 法第 28 条乃至 39 条に規定されている。各構成国が指定組織を設置して認証機関の要件を決定する。さらに、高リスク AI システムに対する第三者適合性評価を EU 市場に上市するにあたって実施する義務があり、適合性評価の後、CE マークに指定組織の識別番号を表示する仕組みとなっている。

なお、整合化された技術標準が存在する場合、その基準を満たしていることを要件適合と判断する。標準が存在しない場合、欧州委員会が共通仕様を策定し、それを満たしていれば要件適合とする仕組みとなっている。

2. 4. 標準化要請の法的枠組み

欧州委員会は、標準化規則 (Regulation (EU) No 1025/2012¹³) に基づき、欧州標準化機

¹² Oystein Endal, Andrea Vetric, Sidsel Nag, Nick Malter & Daylan Araz, *Providers of General-Purpose AI Models — What We Know About Who Will Qualify, EU Artificial Intelligence Act* (25 Apr. 2025), <<https://artificialintelligenceact.eu/providers-of-general-purpose-ai-models-what-we-know-about-who-will-qualify/>>では、世界で約 11 社がこの閾値を超えるモデルを提供しているとし、Robi Rahman, Lovis Heindrich, David Owen & Luke Emberson, *Over 30 AI models have been trained at the scale of GPT-4, Epoch AI Data Insights* (30 Jan. 2025, last updated 6 Jun. 2025), <<https://epoch.ai/data-insights/models-over-1e25-flop>>によれば、 10^{25} FLOPs を超える訓練計算量で開発された AI モデルを 33 件特定している。

¹³ 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

関（CEN、CENELEC、ETSI¹⁴）に対して標準化要請（Standardisation Request: SR）を
発出している。2023年5月22日に採択されたC(2023)3215¹⁵であり、その後AI法の最終
文言に整合するよう2025年6月23日にSRが改定¹⁶されている。

この標準化要請を受けて、CEN/CENELEC JTC 21（合同技術委員会 21）が欧州規格
（EN）を策定し、当該規格に基づいて「適合推定（presumption of conformity）」を得る
ことができる（AI法第40条）。なお、標準が受諾されない場合や遅延・不十分な場合に
は、欧州委員会は同法第41条に基づき共通仕様（Common Specifications, CS）を採択す
る権限を有している。

2. 5. 標準化要請の要求事項

標準化要請には10項目の要求事項が含まれており、①AIシステムのリスク管理システ
ム、②データ及びデータガバナンス、③ログ機能による記録管理、④ユーザへの透明性と情
報提供、⑤人間による監督、⑥AIシステムの精度仕様、⑦AIシステムの堅牢性に関する仕
様、⑧AIシステムのサイバーセキュリティ仕様、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、⑩AIシステムの適合性評価とい
う、AIシステムの信頼性確保に必要な要素が網羅されている。

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, OJ L 316, 14.11.2012,
pp. 12–33.

¹⁴ 検討実務の詳細については、Carlos Lopez Rodriguez, *Status of AI related EU Policy Initiatives*, presentation at ETSI AI Conference 2025, February 10, 2025,
<https://docbox.etsi.org/Workshop/2025/02_AICONFERENCE/SESSION02/EC_DGCNECT_CARLOS_LOPEZ_RODRIGUEZ.pdf>.

¹⁵ European Commission, Commission Implementing Decision of 22.5.2023 on a
standardisation request to the European Committee for Standardisation and the
European Committee for Electrotechnical Standardisation in support of Union policy
on artificial intelligence, C(2023) 3215 final (22 May
2023)<[https://ec.europa.eu/transparency/documents-
register/detail?ref=C\(2023\)3215&lang=en](https://ec.europa.eu/transparency/documents-register/detail?ref=C(2023)3215&lang=en)>.

¹⁶ European Commission, Commission Implementing Decision of 23.6.2025 on a
standardisation request to the European Committee for Standardisation and the
European Committee for Electrotechnical Standardisation as regards high-risk AI-
systems in support of Regulation (EU) 2024/1689 of the European Parliament and of
the Council and repealing Implementing Decision C(2023)3215, C(2025) 3871 final (23
June 2025), <[https://ec.europa.eu/transparency/documents-
register/detail?ref=C\(2025\)3871&lan](https://ec.europa.eu/transparency/documents-register/detail?ref=C(2025)3871&lan)>.

2. 6. 標準化作業の現状

標準化要請の実施期限は当初 2024 年 1 月 31 日であったが、4 月 30 日、さらに 8 月 31 日へと延長され、主要規格の整備に向けた作業は全体的に遅延している。

特に注目すべきは、JTC 21/WG 4 (Foundational and Societal Aspects) が開発を進める prEN 18229 (AI Trustworthiness Framework) である。この規格は、記録保持 (record keeping)、透明性 (transparency)、人間による監督 (human oversight)、精度 (accuracy)、堅牢性 (robustness) といった複数の標準化要請項目を包括する欧州整合規格として設計されている。しかし、対象項目が多岐にわたるため、分冊化 (分割) も議論されており、時間内の欧州規格完成を最優先とする状況となっている。

この遅延により、2026 年 8 月の高リスク AI に係る義務の適用開始が 2027 年 12 月 2 日へ最大 16 ヶ月延期予定ではあるものの、AI 法の適用において整合規格の整備が完了していない状況では、共通仕様で対応せざるを得ない可能性がある。整合規格なしに高リスク AI に関する義務を適用する場合、AI 法第 41 条に基づく「要件適合推定」という曖昧な手続きに依存することになり、法的不確実性が増大することになりかねない。

2. 7. ブリュッセル効果の変容と EU 規制政策の転換

EU が提案した規制が実質的に世界的なルール形成に影響を及ぼすことを「ブリュッセル効果」と呼ぶ見解¹⁷がある。ブリュッセル効果は、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 つの条件を満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をいう。EU の新たな AI 規制は、AI の研究開発及び社会実装に向けて文字通りのブリュッセル効果を発揮する分野になると考えられてきた。ところが、EU からの英国の離脱 (ブレグジット) 以降、ブリュッセル効果の陰りとも言える状況が見受けられる。

その傾向が顕著な例として、厳格な内燃機関車販売禁止政策の方針転換があげられる。EV 政策については、当初 2035 年までに新車の CO₂ 排出を 100%削減 (内燃機関車販売の事実上の禁止) とする目標が設定されていた。ところが、2023 年 3 月にはドイツの反対¹⁸により合成燃料 (e-fuel) 使用車が例外として認められるなど、厳格な方針に一部修正が加えられた。しかし、2025 年 12 月 16 日、欧州委員会は「Automotive Package¹⁹」を発表し、2035 年の内燃機関車販売禁止を事実上撤回した。新たな目標は、2035 年までに CO₂

¹⁷ Anu Bradford, *The Brussels Effect: How the European Union Rules the World*, Oxford University Press, 2020 (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 年)。

¹⁸ Kate Abnett, *EU-German deal to map path for e-fuel cars after 2035*, Reuters, March 27, 2023, <<https://www.reuters.com/business/autos-transportation/eu-german-deal-maps-legal-path-e-fuel-cars-after-2035-document-2023-03-27/>>.

¹⁹ European Commission, *Communication from the Commission: Automotive Package (16 December 2025)*, <https://single-market-economy.ec.europa.eu/publications/automotive-package-documents_en>.

排出を 90%削減（2021 年比）とするものであり、残り 10%については EU 製の低炭素鉄鋼の使用（最大 7%）や e-fuel・バイオ燃料の使用（最大 3%）により相殺することが認められる。これにより、プラグインハイブリッド（PHEV：外部充電可能なハイブリッド車）、マイルドハイブリッド（小型モーターで内燃機関を補助する車両）、レンジエクステンダー（航続距離延長用の発電機を搭載した電気自動車）、及び内燃機関車（ガソリン・ディーゼルエンジン車）が 2035 年以降も販売可能となった。

環境政策においても、2024 年初頭にフランス、ドイツ、オランダ等で農家による大規模抗議が発生し、欧州グリーンディールの農業関連規制が緩和²⁰されている。

その他、森林破壊規制法（EUDR）の施行延期²¹、中国製 EV への関税導入を巡る EU 加盟国間の対立²²など枚挙に暇がなく厳格な規制を維持することの困難さが顕在化している。

AI 法についても、フランス等からイノベーション阻害への懸念²³が示され、整合規格の策定遅延²⁴が生じている。EU の厳格な規制が域内外からの圧力により後退する傾向は、次

²⁰ European Commission, *EU actions to address farmers' concerns*, last modified March 15, 2024, <https://agriculture.ec.europa.eu/overview-vision-agriculture-food/eu-actions-address-farmers-concerns_en>.

²¹ Council of the European Union, "Deforestation: Council signs off targeted revision to simplify and postpone the regulation," press release, December 18, 2025, <<https://www.consilium.europa.eu/en/press/press-releases/2025/12/18/deforestation-council-signs-off-targeted-revision-to-simplify-and-postpone-the-regulation/>>.

²² European Commission, *EU Commission imposes countervailing duties on imports of battery electric vehicles (BEVs) from China*, October 29, 2024, <<https://trade.ec.europa.eu/access-to-markets/en/news/eu-commission-imposes-countervailing-duties-imports-battery-electric-vehicles-bevs-china>>.

²³ IT, FR and DE, *An innovation-friendly approach based on European values for the AI Act: Joint Non-paper by IT, FR and DE*, November 2023, <https://logistic-natives.com/wp-content/uploads/2023/11/9aaf68af_2128_4aa6_b8d6_a61212e5fd46_France_German_231122_140916.pdf>では、イタリア・フランス・ドイツの共同「ノンペーパー」として規制が欧州のイノベーション能力を阻害しないよう主張する趣旨が記載されている。Charles-Edouard Amaelle, *France keeps up its pressure on the EU's AI Act, despite mounting criticism*, Le Monde, January 27, 2024, <https://www.lemonde.fr/en/economy/article/2024/01/27/france-keeps-up-its-pressure-on-the-eu-s-ai-act-despite-mounting-criticism_6471038_19.html>では、フランスが AI 法の一部規定について「イノベーションを阻害し、欧州 AI スタートアップの妨げになる」と繰り返し主張してきた旨が明記されており、フランス経済省の見解として、規制が「イノベーションやテック・エコシステムの発展を損なってはならない」とするコメントが引用されている。

²⁴ CEN-CENELEC, *European standardization organizations move forward on AI standardization*, October 23, 2025, <<https://www.cencenelec.eu/news->

のオムニバス・パッケージに見られる通り、AI分野においても例外ではない可能性がある。

2. 8. オムニバス・パッケージ Omnibus VII (デジタル)

EU 規制政策の転換の動きを体系的に示すものが、2025年11月19日に欧州委員会が公表した10分野のオムニバス・パッケージである。複数の関連法令の相互作用・重複を横断的に調整し、実施・執行や事業者負担を簡素化するため、複数法令を同時に改正する「Omnibus (包括改正) 提案 (packages)」として構成される。7分野目の Omnibus VII (デジタル) は、「デジタル・オムニバス²⁵ (Digital Omnibus (COM(2025) 837))」及び「デジタル・オムニバス AI²⁶ (Digital Omnibus on AI (COM(2025) 836))」から構成される。

3. 日本の AI 制度標準化

3. 1. 政策文書における標準化の位置づけ

「AI 推進法²⁷」が制定され、AI 制度標準化に向けた政策としては、「イノベーション戦略 2024²⁸」において AI 法の制度設計における上位戦略文書として標準化を政策の基盤に据えることが示されている。

2024年4月に公表された「AI 事業者ガイドライン²⁹」は、AI システムの開発・提供・利用に関する包括的な指針を示している。このガイドラインは、国際的に見ても精緻な内容となっており、前述した EU の一般目的 AI 実務準則と比較しても遜色ない水準にある。これらガイドラインによって培ってきた取り組みを礎として、AI システムの管理・運用のために構築するマネジメントシステムを第三者認証により担保する仕組みとして、次節で紹介

[events/news/2025/brief-news/2025-10-23-ai-standardization/](https://www.soumu.go.jp/main_content/news/2025/brief-news/2025-10-23-ai-standardization/)>.

²⁵ Proposal for a Regulation of the European Parliament and of the Council amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework, and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868, and Directive (EU) 2019/1024 (Digital Omnibus) COM/2025/837.

²⁶ Proposal for a Regulation of the European Parliament and of the Council amending Regulations (EU) 2024/1689 and (EU) 2018/1139 as regards the simplification of the implementation of harmonised rules on artificial intelligence (Digital Omnibus on AI (COM(2025) 836 final).

²⁷ 人工知能関連技術の研究開発及び活用の推進に関する法律 (令和7年法律第53号)。

²⁸ 統合イノベーション戦略推進会議「統合イノベーション戦略 2024」(令和6年6月4日閣議決定)。AI 戦略会議が取りまとめた「AI に関する暫定的な論点整理 (令和5年5月26日)」では、日本の強みである「ものづくり」の視点から AI 統治を捉えることの重要性が指摘されている。

²⁹ 総務省・経済産業省「AI 事業者ガイドライン (第1.0版)」(令和6年4月19日)。令和6年11月22日に第1.01版、令和7年3月28日に第1.1版が公表されている。

介する JIS Q 42001 が機能することで、自主的な取り組みと客観的な評価の両立が可能となる。

AI セーフティ・インスティテュート (AISI) が 2024 年 2 月に設立³⁰され、AI の安全性に関する評価基準や手法の検討も開始されている。AISI は、国際的な AI セーフティ・ネットワークの一翼を担い、米国、英国、EU 等との連携を進めている。標準化活動と AISI の取り組みは相互補完的な関係にあり、技術的な安全性評価と組織的なマネジメントシステムの両面から AI 統治の強化が図られている。

3. 2. JIS Q 42001 の制定とその意義

JIS Q 42001³¹ (AI マネジメントシステム規格) は、ISO/IEC 42001 を国内採用する形で策定され、2025 年 8 月 20 日に公示された。これにより、国際的な相互認証が可能となり、AI マネジメントシステムを構築した日本企業の海外展開を支援するとともに、国内における AI の組織的管理体制の整備を促進する基盤が整った。

EU の AI 法に基づく整合規格の策定が当初予定よりも遅れる中、JIS Q 42001 の制定は国際的な AI 統治の議論において日本の発言力を高める効果も期待できる。AI 規制分野における「規制後進国」というイメージを払拭し、「標準先進国」としての地位を確立する好機であることを認識すべきである。

3. 3. 規格の構造

JIS Q 42001 の特徴は、マネジメントシステム規格の共通規格 (旧称 Annex SL、現在は Harmonized Structure) に完全準拠していることである。この構造により、既存のマネジメントシステムとの統合が極めて容易になっている。

共通規格は、全ての ISO マネジメントシステム規格で共通の章立て、共通の用語、共通の要求事項 (いわゆる「ブルーテキスト」) を使用する。例えば、第 4 章「組織の状況」、第 5 章「リーダーシップ」、第 6 章「計画」、第 7 章「支援」、第 8 章「運用」、第 9 章「パフォーマンス評価」、第 10 章「改善」という構造は、全ての規格で共通である。ブルーテキスト部分は変更不可の共通要求事項であり、各規格の固有要求事項のみが追加される。JIS Q 42001 では、AI システム特有の要求事項として、AI リスクアセスメント、AI システム影響評価、透明性の確保、人間による監視等が追加されている。

他のマネジメントシステム規格と同様の整合性を確保しているため、これまでマネジメントシステムに取り組んできた組織にとって理解しやすい構造となっている。管理目的、管理策、実施の手引きは附属書 A、B、C、D に記載されている。

企業が既に既存のマネジメントシステム規格に基づく認証を取得している場合、JIS Q

³⁰ 経済産業省「AI セーフティ・インスティテュート (AISI) を設立しました」(2024 年 2 月 14 日) <<https://www.meti.go.jp/press/2023/02/20240214002/20240214002.html>>。

³¹ JIS Q 42001:2025「情報技術—人工知能—マネジメントシステム」。解説は、高村博紀・杉村領一・原田要之助・新保史生・畔津布枝・坂本静生・保木野昌稔・若井秀子・江川尚志『ISO/IEC 42001:2023 (JIS Q 42001:2025) 情報技術—人工知能—マネジメントシステム要求事項の解説』日本規格協会 (2025 年)。

42001 の追加は差分管理のみで対応可能である。この統合により、マネジメントレビューや内部監査を統合実施できるため経営資源の効率的活用が可能となり、文書管理システムを共通化できるため管理負荷が軽減されるとともに、マネジメントシステムに準拠した対応への理解も容易であり導入期間短縮も期待されるところである。

3. 4. 認定・認証体制の整備

AI マネジメントシステム (AIMS) 適合性評価制度は、認定機関である ISMS-AC (一般社団法人情報マネジメントシステム認定センター) が 2025 年 7 月に認証機関の「認定」業務を開始³²し、2026 年 1 月には 2 つの「認証機関」が認定³³を受けている。本制度では、認定機関 (ISMS-AC) が認証機関を認定・監督し、認定を受けた認証機関が申請事業者の AIMS を審査・認証するという明確な役割分担により、国際的な相互承認も可能である。

3. 5. AI マネジメントシステム適合性評価制度の今後の展開

JIS Q 42001 の制定とともに、本稿執筆時点で JIS Q 42005 及び 42006 の二つの規格整備に向けた検討が行われている。

JIS Q 42005 (ISO/IEC 42005:2025)は、AI システムの影響評価 (インパクトアセスメント) に関する規格である。AI システムが社会、個人、環境に与える影響を体系的に評価する手法を提供する。これは、EU の AI 法が要求する影響評価にも対応可能な内容となっている。

JIS Q 42006 (ISO/IEC 42006:2025)は、AI マネジメントシステムの審査・認証機関に関する要求事項を定める規格である。AI 特有の技術的知識と審査スキルを認証機関の審査員に要求するものである。

4. 国際標準における日本の貢献

4. 1. ISO/IEC における日本の活動

日本は国際標準化に多大な貢献をしている。ISO/IEC JTC 1 (Joint Technical Committee 1) における日本の貢献は、一般に認識されている以上に大きい。AI に関する国際標準の検討は、2017 年 10 月の ISO/IEC JTC 1 総会 (ウラジオストク) において ISO/IEC JTC 1/SC 42³⁴の設置が決議されたことに始まる。翌 2018 年 4 月に北京で第 1 回会合が開催され活動が本格化した。

³² 一般社団法人情報マネジメントシステム認定センター (ISMS-AC) 「AI マネジメントシステム 認証制度の開始について」 (令和 7 年 7 月 8 日) <<https://isms.jp/topics/news/20250131.html>>。

³³ 一般社団法人情報マネジメントシステム認定センター (ISMS-AC) 「AI マネジメントシステム (AIMS) 認証機関の初認定について」 (令和 8 年 1 月 14 日) <<https://isms.jp/topics/news/20260114.html>>。

³⁴ ISO/IEC JTC 1/SC 42, Artificial intelligence, International Organization for Standardization, (accessed January 28), 2026, <<https://www.iso.org/committee/6794475.html>>。

WG 1 (基盤標準)、WG 2 (データ)、WG 3 (信頼性)、WG 4 (ユースケース及びアプリケーション)、WG 5 (AI システムの計算的アプローチと特性)の 5 つの WG が設置されている。ならびに、複数の ISO/IEC 技術委員会 (TC/SC) が共通課題に対応するために合同で設置する作業部会として、JWG 2 (AI ベースシステムのテスト)、JWG 3 (AI 対応ヘルスインフォマティクス)、JWG 4 (機能安全と AI システム)、JWG 5 (自然言語処理)、JWG 6 (AI システムの適合性評価)、JWG 7 (人工知能 (金融サービス)) の 6 つの JWG が設置されている。

これらのワーキンググループのうち、WG 4 (ユースケース及びアプリケーション) および JWG 3 (AI 対応ヘルスインフォマティクス) では日本がコンビナー (議長) を務めており、WG 4 では実用的な AI 応用の標準化において日本の産業界の経験を活かした活動を主導している。さらに、AI システムの性能評価方法と信頼性確保の枠組みという AI 標準化の中核的な領域を扱う WG 2、WG 3 及び JWG 3 においても、日本はエディターの輩出やプロジェクトへの積極的な貢献を通じて重要な役割を果たしている。この事実は、国際標準化における日本の専門性、技術力と組織力の高さを示している。

標準化における主導的立場は、標準の技術的内容に日本の知見を反映させる実質的な機会となっている。特に、品質管理やリスク管理の分野で培われた日本の産業界の経験が国際標準に組み込まれることの意義は大きい。

4. 2. 主要な国際規格の体系

ISO/IEC JTC1/SC 42 は、AI システムのライフサイクル全体をカバーする包括的な国際標準体系³⁵を構成している。基盤・用語分野では ISO/IEC 22989 (AI 概念及び用語)、

³⁵ 本稿執筆時点における国際標準の整備状況は以下の通り。

【基盤・用語】 [発行] ISO/IEC 22989:2022 (AI 概念と用語)、ISO/IEC 23053:2022 (ML 枠組み)、ISO/IEC TR 17903:2024 (ML コンピューティング装置)。 [開発中] ISO/IEC 22989-2 (医療用語)、ISO/IEC 22989/Amd 1・23053/Amd 1 (生成 AI 追加)、ISO/IEC 12792 (透明性タクソノミー) 等。

【データ品質・ライフサイクル】 [発行] ISO/IEC 5259-1~5:2024-25 (データ品質シリーズ)、ISO/IEC 8183:2023 (データライフサイクル枠組み)。 [開発中] ISO/IEC 25590 (生成 AI 出力データ品質) 等。

【AI システムライフサイクル】 [発行] ISO/IEC 5338:2023 (ライフサイクルプロセス)、ISO/IEC 5339:2024 (AI 応用指針)、ISO/IEC 5392:2024 (知識工学参照アーキテクチャ)。

【リスク管理・安全・ガバナンス】 [発行] ISO/IEC 23894:2023 (リスク管理指針)、ISO/IEC TR 5469:2024 (機能安全と AI)、ISO/IEC 38507:2022 (AI ガバナンス)、ISO/IEC 42001:2023 (AI マネジメントシステム)、ISO/IEC 42005:2025 (影響評価)、ISO/IEC 42006:2025 (AIMS 認証機関要件)。 [開発中] ISO/IEC 42003 (42001 実装指針)、ISO/IEC 42007 (適合性評価スキーム枠組み)、ISO/IEC TS 25568 (生成 AI リスク対応指針)。

【公平性・倫理・社会的側面】 [発行] ISO/IEC TR 24027:2021 (バイアス)、ISO/IEC

ISO/IEC 23053（機械学習を用いる AI システムの枠組み）等が策定されている。データ品質分野では ISO/IEC 5259 シリーズ（分析及び機械学習のためのデータ品質）、ISO/IEC 8183（データライフサイクル枠組み）が重要である。

リスク管理・ガバナンス分野は特に充実しており、ISO/IEC 23894（AI リスク管理に関する手引）、ISO/IEC 42001（AI マネジメントシステム）、ISO/IEC 42005（AI システム影響評価）、ISO/IEC TR 5469（機能安全と AI）等、複数の標準が相互補完的に策定されている。公平性・倫理分野では ISO/IEC TR 24027（バイアス）、ISO/IEC TR 24368（倫理的・社会的懸念の概観）、堅牢性・信頼性分野では ISO/IEC 24029 シリーズ（ニューラルネットワークの堅牢性評価）、品質評価分野では ISO/IEC 25059³⁶（AI 品質モデル）が発行されている。説明可能性分野では ISO/IEC TS 6254（ML モデル及び AI システムの説明可能性・解釈性の目標とアプローチ）が発行され、ヒューマンオーバーサイト分野では ISO/IEC 42105（AI システムの人間による監督に関する手引）等が検討されている。

TS 12791:2024（バイアス処理）、ISO/IEC TR 24368:2022（倫理的・社会的懸念）、ISO/IEC TR 21221:2025（有益な AI システム）、ISO/IEC TR 20226:2025（環境持続可能性）。【ロバストネス・信頼性】〔発行〕ISO/IEC TR 24028:2020（信頼性概観）、ISO/IEC TR 24029-1:2021（堅牢性評価概観）、ISO/IEC 24029-2:2023（形式手法）。

〔開発中〕ISO/IEC 24029-3（統計的方法）等。

【品質評価】〔発行〕ISO/IEC TS 25058:2024（AI 品質評価指針）、ISO/IEC 25059:2023（AI 品質モデル）。

【説明可能性・性能測定】〔発行〕ISO/IEC TS 4213:2022（分類性能評価）、ISO/IEC TS 6254:2025（説明可能性・解釈性）、ISO/IEC TS 8200:2024（AI 制御可能性）。〔開発中〕ISO/IEC TR 42103（合成データ概観）等。

【テスト・監査】〔開発中〕ISO/IEC TS 42119 シリーズ（AI テスト）、ISO/IEC 25870（インシデント報告）、ISO/IEC 24970（AI システムログ）。

【機能安全】〔発行〕ISO/IEC TR 5469:2024。〔開発中〕ISO/IEC TS 22440 シリーズ（機能安全要件・指針・適用例）。

【人間中心・ヒューマンマシンチーミング】〔開発中〕ISO/IEC 25589（チーミング枠組み）、ISO/IEC 42105（人間による監督）、ISO/IEC TR 42109（ユースケース）。

【応用分野・ユースケース】〔発行〕ISO/IEC TR 24030:2024（AI ユースケース）、ISO/IEC TR 24372:2021（計算的アプローチ概観）。〔開発中〕ISO/IEC TR 18988（医療 AI 応用）等。

【ビッグデータ】〔発行〕ISO/IEC 20546:2019（用語）、ISO/IEC TR 20547 シリーズ（参照アーキテクチャ）、ISO/IEC 24668:2022（分析プロセスマネジメント）。

³⁶ ISO/IEC 25059 は ISO/IEC JTC 1/SC 7（ソフトウェア及びシステム技術）の SQuaRE（システム及びソフトウェア製品の品質要求及び評価）シリーズの一部として策定された標準であり、厳密には SC42 単独の成果物ではなく、SC7 との連携により策定されたものである。

5. 統合認証とサイバネティック・アバター(CA)

5. 1. 仮想化実世界の進展とサイバネティック・アバターに関する研究

AI と人間の共生が加速度的に進む中で、物理的世界とデジタル世界の境界は一層曖昧なものとなり、「仮想化実世界³⁷⁾」の進展が見込まれる。「サイバネティック・アバター（以下、「CA」という。）」は、この融合の象徴的な存在である。CA の社会実装においては、利用者認証（操作者の本人確認）、継続認証（利用中の本人性の継続的確認）、CA 公証（認証された公認 CA としての状態確保）という認証ステップが必要となる。

「統合認証（Integrated Certification³⁸⁾」は、これら CA 固有の認証ステップに加え、CA を構成するロボット等の製品認証と AI システムのマネジメントシステム認証を包括的に扱う枠組みであり、物理的存在とデジタル存在の境界を越えて活動する CA の信頼性を担保するための重要なツールとなる。人間とアバターの関係性を法的に位置づけ、その信頼性を担保する仕組みの構築とともに、統合認証の実現はメタバースやデジタルツインが社会基盤となる時代において不可欠の課題である。

筆者は、JST ムーンショット型研究開発事業目標 1「2050 年までに、人が身体、脳、空間、時間の制約から解放された社会を実現」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（JPMJMS2215）」のプロジェクトマネジャーを務めている（以下、「新保プロジェクト」という）。新保プロジェクトでは、CA の社会実装に向けて、その安全性と信頼性を担保するための制度的枠組みの研究を進めており、統合認証はその中核的課題の一つである。

5. 2. 統合認証の必要性

AI 規正論 2.0 の今後の目標は「統合認証」の実現に向けた理論的基礎を提供することにある。新保プロジェクトでは 2022 年から統合認証の研究³⁹⁾を進めている。

統合認証の有用性は、CA の認証に係る研究を通じてその実用性を認識⁴⁰⁾するとともに、

³⁷⁾ 石黒浩「アバターによる仮想化実世界の倫理問題」人工知能, Vol. 36, No. 5, pp. 558-563 (2021).

³⁸⁾ 「統合認証」の英訳について、利用者の「本人確認」を意味する場合の「認証」は「Authentication」である。ムーンショット研究開発プロジェクトにおいて研究を進めている「統合認証」とは、単なる本人確認にとどまらず、ロボットなどの製品及びそのロボットに組み込まれる AI システムの管理（マネジメント）双方の認証を包含し、有体物 CA と無体物 CA に適用可能な統一的認証フレームワークを指す。仮想化実世界における有体物（物理的なロボット）と無体物（バーチャル・アバターなど）、製品とマネジメントシステムの双方について統合的な認証を実現することを最終的な目的としていることから、基準・規格への適合を第三者が認証することを意味する「Certification」にあたることから「Integrated Certification」と訳している。

³⁹⁾ 統合認証の実現に向けた制度設計の提案については、新保史生「CA の社会実装に向けた適合性評価制度の構築」君嶋祐子＝田中浩之＝麻生典編『サイバネティック・アバター(CA)と法』（弘文堂、2025 年）。

⁴⁰⁾ 統合認証検討の前提としてのサイバネティック・アバター(CA)の分類と認証について

マネジメントシステムを構築する際に複数規格に対応する負担を減らし、効率的な審査も可能にすることが期待される。

昨今のセキュリティリスクの質的变化においても、統合認証の重要性はさらに高まることが想定される。現在のセキュリティ・インシデントの多くは、システム停止やデータ漏洩といった、主に経済的損失や信用失墜にとどまっている。ランサムウェア被害の例では、製品の発注・受注・出荷が停止し、事業継続に重大な影響を与えたが、直接的な人的被害は発生していない。しかし、ロボットやAIシステムが日常生活に深く浸透した社会では、セキュリティ・インシデントの影響は質的に異なる。街中を歩くサービスロボットがハッキングされた場合、物理的な危害が直接人間に及ぶ。自動運転車の大規模ハッキングでは、多数の死傷者が発生する可能性がある。

例えば、「全ての自動運転車が一斉に特定の場所に向かう」といったシナリオは、もはやSFではない。都内を走行するすべての自動運転の車両が首相官邸その他の政府の重要施設に突入しようとしているといった事態は、国家安全保障上の脅威となる。このような事態を防ぐためには、製品レベルの安全性とシステムレベルのセキュリティを統合的に管理する必要がある。

統合認証のもう一つの重要な側面は責任の明確化である。現在の分離された認証体系では、インシデント発生時に「マネジメントシステムは適切だったが製品に欠陥があった」または「製品は安全だったがマネジメントが不適切だった」といった責任の押し付け合いが発生する可能性がある。統合認証により、組織は製品とマネジメントシステムの両方に対して一体的な責任を負うことになる。

付言しておきたい。CAは、物理的な存在（有体物）とデジタルな存在（無体物）の境界を越えて活動する新しいタイプのシステムであり、その分類は多面的であることから、認証の枠組みも従来とは異なるアプローチが必要となる。有体物CAはロボット型のアバターであり、無体物CAはデジタル空間上のアバター（VTuber等）である。利用場面には、①身代わりとして実在の人物の代理、②物故者として亡くなった方の再現、③非実在としてキャラクター等の表現がある。動作の形態は、①遠隔操作、②プログラムによる自動処理、③自律的動作という3つのモードがある。

多様なCA利用環境に対応するため、認証には3つのステップが必要となる。第一に利用者認証であり、記憶（パスワード等）、有体物（ICカード等）、バイオメトリクス（生体認証）による本人確認である。第二に継続認証であり、CA利用中における継続的な本人性の確認である。第三にCA公証であり、認証された公認CAとしての状態を確保する仕組みである。

これらのステップを経て、初めてCAを安全かつ安心して使用できる環境が整う。CAの統合認証においては、AIシステムとしての安全性・信頼性の認証に加え、CA固有の存在証明や操作者認証を包括的に扱う枠組みとしてその実用性を実証できる可能性がある。

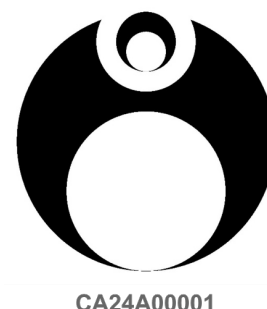
5. 3. CA 認証マークの開発

CA の社会実装においては、利用者がその CA が認証され信頼できる存在であることを容易に確認できる仕組みが不可欠である。新保プロジェクトでは、CA の信頼性を可視化し、利用者が安心して CA を利用できる環境を整備することを目的として、「CA 認証マーク (Certified Avatar Mark)」を開発している。

CA 認証マークのデザインは、大阪大学大学院基礎工学研究科の石黒浩教授によるものである。マークには、偽造防止と認証情報の確実な伝達を実現するため、電子透かし、マイクロ文字、イルミグラム (ホログラム) の技術を実装している。マークの画像データ内に、当該 CA の認証情報 (認証番号、認証日時、認証対象や組織等の管理情報等) が人間の目には見えない形で電子透かしとして埋め込まれており、利用者はスマートフォンでマークを撮影することで、当該 CA の認証情報を確認することができる。

CA 認証マークは、2025 年大阪・関西万博において会場内で稼働するロボットに実装され、来場者が実際に CA 認証システムの動作を体験できる形で社会実装の第一歩を踏み出した。なお、CA 認証システムの概要を説明する動画は YouTube⁴¹で公開されている。

【CA 認証マーク】



5. 4. 統合認証の国際的展開

「統合認証」は、適合性評価に関する検討を行う ISO/IEC JTC 1/SC 42 と ISO/CASCO の合同作業部会である JWG 6 (AI システムの適合性評価スキーム) において検討が行われる予定である。

統合認証の基盤として、2012 年以降、主要な ISO マネジメントシステム規格は、附属書 SL の High Level Structure (HLS) (現在の Harmonized Structure (HS)) に基づく同一構造・共通用語を採用しており、異なる規格間での統合運用が容易になっている。

本稿執筆時点で、統合認証そのものを直接扱う ISO/CASCO 規格は存在しない⁴²。しか

⁴¹ 新保史生「サイバネティック・アバター認証システムについて (動画)」 <<https://www.youtube.com/watch?v=8gzbcu5eY8M>>。

⁴² 統合認証に直接対応する ISO 規格は存在しないが、関連する既存規格・ガイダンスとして以下がある。【附属書 SL (高位構造: HLS)】ISO マネジメントシステム規格の共通枠組み。ISO 9001、ISO 14001、ISO 45001 等の主要規格は同一の章立て・用語を採用して、異なる規格要求事項の統合運用を容易にする基盤。【ISO/IEC 17021-1】マネジメントシステム認証機関に対する要求事項を定める CASCO 規格。認証機関が複数のマネジメントシステムを一括して審査・認証する場合も本規格に基づく認定が必要となる。派生文書として各分野別の能力要求事項 (17021-2 環境、-3 品質等) がある。【IAF MD 11】国際認定フォーラム (IAF) が発行した統合マネジメントシステムの審査に関する文書。正式名称は「IAF Mandatory Document for the Application of ISO/IEC 17021-1 for Audits of Integrated Management Systems」。複数マネジメントシステムの同時審査における手順や留意事項を定め、統合審査を行う場合の審査工数削減ルール (重複部分の考慮等)、統合審査チームの力量、不適合の扱い方等を具体的に規定。最新版は 2023 年改

し、国際認定フォーラム（IAF）が発行した IAF MD 11（統合マネジメントシステムの審査に関する必須文書）が事実上の実務的標準として機能しており、複数のマネジメントシステムの同時審査における手順、審査工数の算定ルール、統合審査チームの力量要件等を規定している。

CASCO の政策会合や年次総会では、「One Audit」（一度の監査で多領域をカバーする考え方）や統合的アプローチが議題⁴³に上っている。2024年4月の第38回 CASCO 総会（ウガンダ・カンパラ）では、デジタル技術やサステナビリティと並び、複数の適合性評価を組み合わせた効率化アプローチが議論され、「ワンストップ型認証サービス」整備の必要性が示唆⁴⁴された。現状では正式な作業部会（WG）や新提案規格（NP）の立上げには至っていないが、予備作業項目（PWI）として各国専門家間で文書草案の非公式検討が進められているとされる。将来的には、IAF MD 11 を ISO/CASCO の正式規格に格上げする形で統合認証ガイドラインを国際標準化するシナリオが想定される。

なお、過去には PAS⁴⁵ 99⁴⁶（BSI 英国規格協会）が統合認証に近い取り組みを行っていたが既に廃止されている。

PAS 99 は、統合マネジメントシステムの先駆的な試みであった。セキュリティ（ISO 27001）、品質（ISO 9001）、環境（ISO 14001）、労働安全衛生（ISO 45001）、事業継続（ISO 22301）、食品安全（ISO 22000）等、複数のマネジメントシステムを統合する画期的な仕様書であった。PAS 99 の革新性は、単に複数の認証を束ねるのではなく、共通のプロセスと統合されたリスク管理により、全体最適化を図った点にある。しかし、10年以上前に公表された時点ではまだそのような規格を活用するには時期尚早であったことや、共通規格（Harmonised Structure）の登場により各規格の統合が容易になったこと等から最終的に廃止となっている。

訂の Issue 3 であり、CASCO の 17021-1 規格と整合しつつ各認定機関に適用が義務付けられている。【ISO 19011】管理システム監査の手引きを提供する規格。複数システムの統合監査を計画・実施する際の監査プログラムマネジメントや監査チーム編成に関する一般指針が含まれる。

⁴³ CASCO Newsletter – December 2023, ISO-CASCO Newsletter, December 2023, 4 pp., Scribd <<https://www.scribd.com/document/709418538/ISO-CASCO-CASCO-Newsletter-December-2023>>.

⁴⁴ IAF Secretariat, *Insights from the 38th Plenary of ISO's Committee for Conformity Assessment*, IAF Outlook, May 31, 2024<<https://iaf.news/2024/05/31/insights-from-the-38th-plenary-of-isos-committee-for-conformity-assessment/>>.

⁴⁵ PAS（Publicly Available Specification）は正式な ISO/IEC 国際規格（IS）ではなく公開仕様書。

⁴⁶ BSI (British Standards Institution), PAS 99:2012 Specification of common management system requirements as a framework for integration, <<https://knowledge.bsigroup.com/products/specification-of-common-management-system-requirements-as-a-framework-for-integration>>（2026年1月28日確認）。

6. おわりに

本稿で提示した AI 規正論 2.0 は、単なる規制対応の枠組みを超えて、信頼できる AI 社会の実現への道筋を示すことを意図するものである。技術革新と安全性の両立、国際競争力と社会的責任の調和、効率性と公平性のバランスなど、AI 統治をめぐる課題は複雑かつ多層的である。これらに対し、国外の規制モデルを単に模倣するのではなく、日本の産業構造や社会的特性を踏まえた独自の解決策を提示することが求められる。

AI 規正論 1.0 で提唱した方向性のうち、産業標準化法に基づく AI マネジメントシステム規格については、JIS Q 42001 の制定により今後の展開を期待できる段階に至っている。日本が培ってきた品質管理と「ものづくり」の伝統を、AI 統治のための規正の枠組みを検討する際に活かし、AI 時代に適応させる道筋も明確になったと言える。今後は、この枠組みを基盤として、AI 統治におけるベストプラクティスを国際社会に発信していくことが、我が国の AI 推進政策において重要な課題となる。

もとより、AI 規正論 2.0 が示す方向性は、未だ検討の途上に過ぎない。本稿の議論が途上であるのみならず、AI 規制の先駆者である EU 自身が AI 法の適用開始スケジュールの延期や規制緩和に向けた一連の「オムニバス提案」を公表し、当初の厳格な規制方針からの軌道修正を迫られている。規制による安全確保と、規制緩和による産業競争力の維持という二律背反に直面する EU の苦悩は、国際的な AI 統治の在り方が未だ定まっていないことを如実に示している。

我が国における今後の AI 推進関連政策においては、このような動向を注視しつつも、日本独自の強みを活かした貢献を模索し続けることが肝要である。信頼できる AI 社会の実現に向けて、本稿がその議論の一助となれば幸いである。

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