国立国会図書館 調査及び立法考査局

Research and Legislative Reference Bureau
National Diet Library

論題 Title	AI 技術と社会―倫理・法・社会的課題（ELSI）と諸外国の AI 規制の動向―
他言語論題 Title in other language	AI Technology and Society: Ethical, Legal, and Social Issues (ELSI) and Global Trends in AI Regulation
著者 / 所属 Author(s)	吉永京子（YOSHINAGA Kyoko）／慶應義塾大学大学院政策・メディア研究科特任准教授
書名 Title of Book	AI と社会のこれからを考える
シリーズ Series	調査資料 2024-4（Research Materials 2024-4）
編集 Editor	国立国会図書館 調査及び立法考査局
発行 Publisher	国立国会図書館
刊行日 Issue Date	2025-3-18
ページ Pages	41-62
ISBN	978-4-87582-937-9
本文の言語 Language	日本語（Japanese）
摘要 Abstract	科学技術に関する調査プロジェクト「AI と社会のこれからを考える」のパネリスト報告

＊ この記事は、調査及び立法考査局内において、国政審議に係る有用性、記述の中立性、客観性及び正確性、論旨の明晰（めいせき）性等の観点からの審査を経たものです。

＊ 本文中の意見にわたる部分は、筆者の個人的見解です。

国立国会図書館
National Diet Library, Japan

AI 技術と社会
―倫理・法・社会的課題（ELSI）と諸外国の AI 規制の動向―　プレゼンテーション資料

国立国会図書館　科学技術に関する調査プロジェクト2024シンポジウム
「AIと社会のこれからを考える」
2024年11月15日

AI技術と社会
―倫理・法・社会的課題（ELSI）と諸外国のAI規制の動向―

吉永 京子
慶應義塾大学大学院　政策・メディア研究科　特任准教授
ジョージタウン大学法科大学院テクノロジー法政策研究所
ノンレジデント・フェロー
東京大学未来ビジョン研究センター　客員研究員

This research was supported by the JST Moonshot R&D project, JPMJMS2215

Keio University Shonan Fujisawa Campus

スライド 1

吉永京子（略歴）

- 東京大学大学院法学政治学研究科修士課程修了（法学修士）
- 2003年４月～2023年6月 株式会社三菱総合研究所
- 2010-11年米国イェール大学ロースクール情報社会プロジェクト(Yale ISP)Visiting Fellow
- 2020年9月～米国ジョージタウン大学ロースクールテクノロジー法政策研究所 Non-Resident Fellow
  （2021年３月～2023年３月、ワシントンD.C.）

Bio: https://www.law.georgetown.edu/tech-institute/people/distinguished-fellows-and-non-resident-fellows/kyoko-yoshinaga/

- 2023年1月～GPAI (Global Partnership on AI) Expert
  Future of Work WG所属、Data Governance WGのco-generated worksプロジェクトの共同リーダー
- 2023年4月～東京大学未来ビジョン研究センター客員研究員
- 2023年10月～慶應義塾大学大学院政策・メディア研究科 特任准教授

その他、
経済産業省「AI事業者ガイドライン」検討会委員、GPAI国内検討会委員
複数の民間企業における「AI倫理委員会」の社外委員
専門：情報通信・メディア・情報セキュリティ関連の法制度・政策、AIガバナンス、AIの法と倫理

2

スライド 2

本日お話すること

1．AI規制の議論の始まり
2．各国のAI規制のアプローチ
3．相互運用性
4．企業に求められること
5．国に求められること
6．汎用型AIの規制

3

スライド 3

1．AI規制の議論の始まり

2016年頃から

＜背景＞

- 2016年3月、ディープラーニングを用いたAIの囲碁プログラムが人間のチャンピオンに勝つ
- リスクが認識される（個人・社会への悪影響）
  - 差別・バイアスの増幅
  - プライバシー侵害
  - AIが仕事を奪う
  - AIの「ブラックボックス問題」

2022年末頃から

生成AIが登場すると・・・さらに、

- 誤情報、ディープフェイク、違法・有害コンテンツ、バイアス（別の意味で）、著作権、プライバシー、その他倫理的問題、働くことや教育にもたらす影響

4

スライド 4

1． AI規制の議論の始まり

- 生成AIに関するプロファイル

NIST-AI-600-1, Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile（2024年7月26日公表）

➢生成AI特有の12のリスクと、開発者が講じ得る400以上の対応策をリスト化した文書。

悪化するリスク
1. CBRN情報または対応能力（「Chemical, Biological, Radiological, and Nuclear（化学、生物、放射性、核）」の略）
2. 捏造(Confabulation)
3. 危険、暴力的、または憎悪を含む内容
4. データプライバシー
5. 環境への影響
6. 有害な偏見や均質化
7. 人間とAIの構成
8. 情報の整合性
9. 情報セキュリティ
10. 知的財産
11. わいせつ、品位を傷つける、または虐待的なコンテンツ
12. バリューチェーンとコンポーネントの統合

5

スライド 5

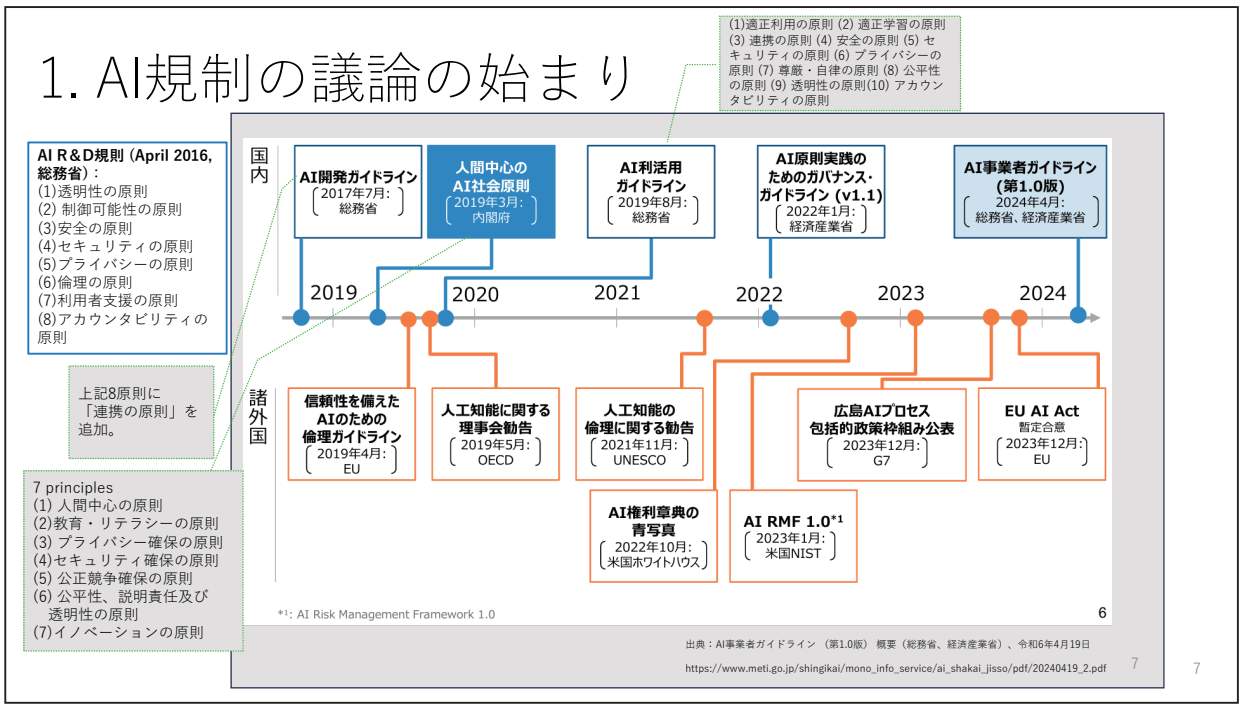
1． AI規制の議論の始まり

ELSI= Ethical, Legal, Social, Issues/ Implications
（倫理的、法的、社会的課題 /意味合い）

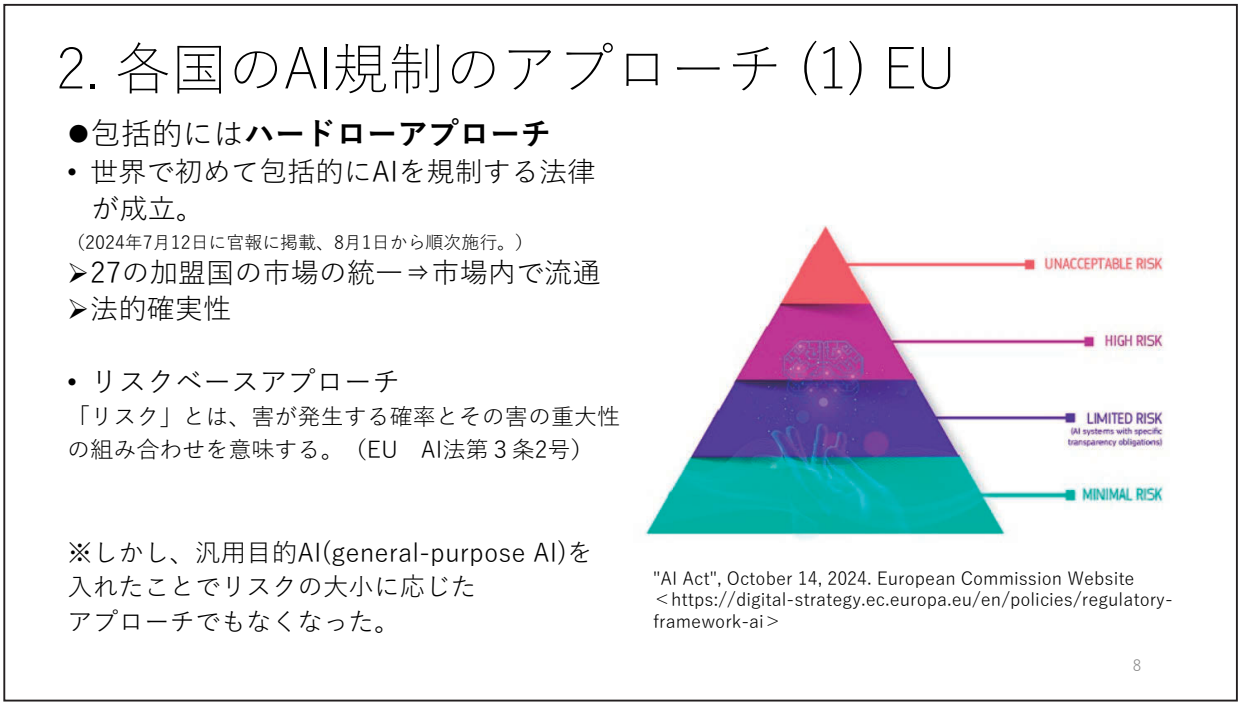
	例
Ethical（倫理的）	バイアス・公平性、透明性、プライバシー、雇用への影響、兵器への使用
Legal （法的）	責任、個人情報保護・プライバシー、知的財産権、差別、法遵守
Social （社会的）	職の喪失と労働力の変革、教育への影響、誤情報の拡散と民主主義への影響、不平等・格差・社会的孤立、デジタル・デバイド、文化への影響・多様性の喪失、ヒューマニティへの影響

6

スライド 6


1. AI規制の議論の始まり
(1)適正利用の原則 (2) 適正学習の原則 (3) 連携の原則 (4) 安全の原則 (5) セキュリティの原則 (6) プライバシーの原則 (7) 尊厳・自律の原則 (8) 公平性の原則 (9) 透明性の原則(10) アカウンタビリティの原則
AI R&D規則 (April 2016, 総務省)：
(1)透明性の原則
(2) 制御可能性の原則
(3)安全の原則
(4)セキュリティの原則
(5)プライバシーの原則
(6)倫理の原則
(7)利用者支援の原則
(8)アカウンタビリティの原則
上記8原則に「連携の原則」を追加。
7 principles
(1) 人間中心の原則
(2)教育・リテラシーの原則
(3) プライバシー確保の原則
(4)セキュリティ確保の原則
(5) 公正競争確保の原則
(6) 公平性、説明責任及び透明性の原則
(7)イノベーションの原則
国内
諸外国
AI開発ガイドライン (2017年7月: 総務省)
人間中心のAI社会原則 (2019年3月: 内閣府)
AI利活用ガイドライン (2019年8月: 総務省)
AI原則実践のためのガバナンス・ガイドライン (v1.1) (2022年1月: 経済産業省)
AI事業者ガイドライン (第1.0版) (2024年4月: 総務省、経済産業省)
2019
2020
2021
2022
2023
2024
信頼性を備えたAIのための倫理ガイドライン (2019年4月: EU)
人工知能に関する理事会勧告 (2019年5月: OECD)
人工知能の倫理に関する勧告 (2021年11月: UNESCO)
広島AIプロセス包括的政策枠組み公表 (2023年12月: G7)
EU AI Act 暫定合意 (2023年12月: EU)
AI権利章典の青写真 (2022年10月: 米国ホワイトハウス)
AI RMF 1.0*1 (2023年1月: 米国NIST)
*1: AI Risk Management Framework 1.0
6
出典：AI事業者ガイドライン （第1.0版） 概要（総務省、経済産業省）、令和6年4月19日
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240419_2.pdf
7
7

スライド 7


2. 各国のAI規制のアプローチ (1) EU
●包括的には**ハードローアプローチ**
• 世界で初めて包括的にAIを規制する法律が成立。
（2024年7月12日に官報に掲載、8月1日から順次施行。）
➢27の加盟国の市場の統一⇒市場内で流通
➢法的確実性
• リスクベースアプローチ
「リスク」とは、害が発生する確率とその害の重大性の組み合わせを意味する。（EU　AI法第３条2号）
※しかし、汎用目的AI(general-purpose AI)を入れたことでリスクの大小に応じたアプローチでもなくなった。
UNACCEPTABLE RISK
HIGH RISK
LIMITED RISK
(AI systems with specific transparency obligations)
MINIMAL RISK
"AI Act", October 14, 2024. European Commission Website
＜https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai＞
8

スライド 8

2. 各国のAI規制のアプローチ (1) EU

規制対象	概要
（1）禁止されたAI行為	サブリミナルな技法や操作的な技法、年齢・障がい等の脆弱性情報を用いて人間の行動・意思決定に悪影響を及ぼすこと、ソーシャルスコアリングを行って人間に害を及ぼすこと、プロファイリング等のみに基づいた犯行予測、インターネット等からスクレイピングした顔認識DBを作成すること、職場や教育機関における自然人の感情を推測すること、生体情報分類システムの使用、公共の場でのリアルタイムの遠隔生体識別システムを法執行目的で使用すること（例外あり）。
（2）ハイリスクAIシステム	ハイリスクAIシステムに求められる要求事項（リスクマネジメントシステムの構築、データガバナンス、技術文書、関係者への情報開示等の透明性、人間による監督等）、プロバイダー・デプロイヤー等の義務（プロバイダーは、例えば、品質管理システム、文書の保管、ログ保存、事前の適合性評価、リスク対応等）。
（3）特定のAIシステム	プロバイダーとデプロイヤーに対する透明性の要求事項。 例：エンドユーザが接しているのはAI（チャットボットやディープフェイク）であることを通知すること。
（4）汎用目的AIモデル (general-purpose AI model; GPAI)	プロバイダーの義務（技術文書・仕様手順書の提供、EU著作権指令を遵守、事前学習に使用されたコンテンツの概要の公開等）。 システミックリスクがあるGPAIプロバイダーの義務（モデル評価、敵対的テストの実施、重大なインシデントの追跡・報告、サイバーセキュリティの確保等）。（市場投入前のR＆D目的のモデルは対象外）

プロバイダー：開発者・提供事業者、デプロイヤー：利活用事業者・ユーザー事業者

9

スライド 9

2. 各国のAI規制のアプローチ (2) 米国

●連邦レベル：**包括的にはソフトローアプローチ＋分野別にハードロー**
●自治体レベル：**ハードロー**もある。

＜連邦＞
- 2022年10月、ホワイトハウスの科学技術政策局（OSTP）
「AI権利章典の構想（青写真）」
The Blueprint for an AI Bill of Rights
- 2023年7月と9月：バイデン政権がAI開発先行企業（7月に7社、9月にさらに8社）を招集し、自主的な取り組みとして安全でセキュアで信頼性のあるAI開発を約束させた（Voluntary AI Commitments)。
- 2023年1月、NIST「**AIリスクマネジメントフレームワーク1.0**」
（日本語訳：https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.jpn.pdf）
- 2023年10月、「**人工知能（AI）の安心、安全で信頼できる開発と利用に関する大統領令**」(Executive Order, E.O.)

＜自治体レベル＞
- 2021年12月、ニューヨーク市議会　ニューヨーク市に住む被雇用者と求職者に対して採用活動におけるAI利用に関する通知（少なくとも10営業日前に通知）、ツールに利用されるデータの種類や出所、雇用者側のデータ保持に関する方針の通知等を規定。

10

スライド 10

2. 各国のAI規制のアプローチ (2) 米国

＜その他の最近の動き＞

- 2024年7月25日、米国国務省（The U.S. Department of State）が “Risk Management Profile for Artificial Intelligence and Human Rights”を公開。（組織（政府、民間企業、市民団体）が国際的な人権に沿う形で設計、開発、利活用、統治するための実践的ガイド） https://www.state.gov/risk-management-profile-for-ai-and-human-rights/ （注 1）

- 2024年7月26日、米国NIST　生成AIに関するプロファイルを公表。
  NIST-AI-600-1, Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile
  ➢生成AI特有の12のリスクと、開発者が講じ得る400以上の対応策をリスト化した文書。
  https://www.nist.gov/artificial-intelligence/executive-order-safe-secure-and-trustworthy-artificial-intelligence-1 （注 2）

- 2024年9月18日、米国と日本がNIST　AI―RMFとAI事業者ガイドラインとのCrosswalk2を同時公開。
  ＜https://airc.nist.gov/AI_RMF_Knowledge_Base/Crosswalks＞
  ＜https://aisi.go.jp/assets/pdf/AISI_Crosswalk2_RMF_GfB_ver1.0.pdf＞

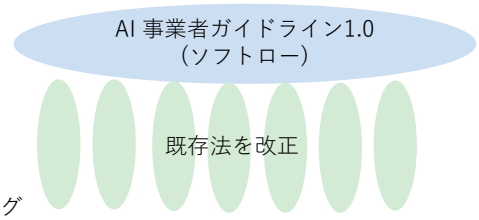
- カリフォルニア州法

11

スライド 11

2. 各国のAI規制のアプローチ (3) 日本

- 包括的には**ソフトローアプローチ**
  「AI事業者ガイドライン1.0」



＜背景＞

➢ 社会課題の解決手段として、AIの活用が期待されていること
　少子高齢化に伴う労働力の低下等

➢ 法律（整備と施行）と技術（スピード、複雑さ）とのタイムラグ

➢ 細やかな行為義務を規定するルールベースの規制はイノベーションを阻害する可能性があること

目標：Society 5.0”-サイバー空間とフィジカル空間を高度に融合させたシステム（CPS：サイバー・フィジカルシステム）による**経済発展**と**社会的課題の解決を両立**する**人間中心の社会の実現**を目指す。

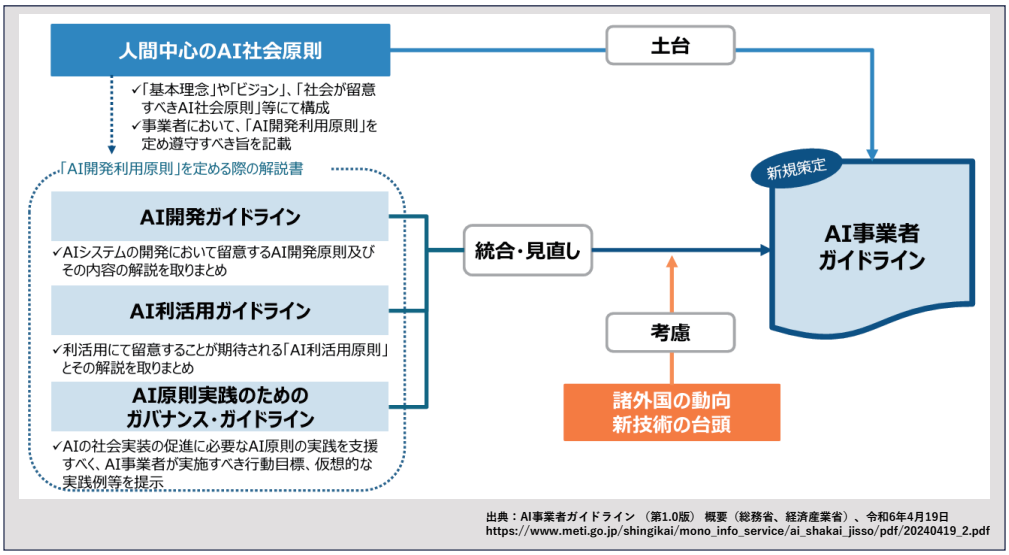
出典：AI事業者ガイドライン （第1.0版） （総務省、経済産業省）、令和6年4月19日
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240419_1.pdf

- **分野別**には必要に応じて**ハードロー**で対応（**既存法の改正**）も。
  例：特定デジタルプラットフォームの透明性及び公正性の向上に関する法律、金融商品取引法
  AIの利活用促進のための改正も。例：道路交通法

12

スライド 12

2. 各国のAI規制のアプローチ (3) 日本



13

スライド 13

2. 各国のAI規制のアプローチ (3) 日本

- **リスクベース・アプローチ**
- **アジャイル・ガバナンス**
- **事業者の自主的な取組の支援**
- **国際的な議論との協調**
- **マルチステークホルダー**(教育・研究機関、一般消費者を含む市民社会、民間企業等)
- **Living Document**

構成

本編: *Why, What*
定義
基本理念、原則
事業者共通の指針、高度なAIシステムに関する事業者に共通の指針、主体別（AI開発者、AI提供者、AI利用者）のBasic Principles
for all AI business actors and for each AI business actor (AI Developers, AI Providers and AI Business Users)

別添: *How*
解説、取組事例、「AI・データの利用に関する契約ガイドライン」を参照する際の留意事項、チェックリスト、主体横断的な仮想事例、海外ガイドライン等の参照先

本編と別添:
総務省: https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02ryutsu20_04000019.html
経済産業省: https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/20240419_report.html

14

スライド 14

2. 各国のAI規制のアプローチ (3) 日本

＜その他の最近の動き＞

- 2024年9月18日：日本、米国がAI事業者ガイドラインとNIST AI—RMFのCrosswalk2を同時公開。

＜https://airc.nist.gov/AI_RMF_Knowledge_Base/Crosswalks＞
＜https://aisi.go.jp/assets/pdf/AISI_Crosswalk2_RMF_GfB_ver1.0.pdf＞

- 2024年9月18日：AISIが「AIセーフティに関する評価観点ガイド（日本語版、英語版）」を公開。

＜https://www.ipa.go.jp/pressrelease/2024/press20240918-2.html＞

- 2024年9月25日：AISIが「AIセーフティに関するレッドチーミング手法ガイド（日本語版、英語版）」を公開。

<https://aisi.go.jp/effort/effort_information/240925/＞

AIセーフティ・インスティテュート（AISI）：https://aisi.go.jp/

15

スライド 15

2. 各国のAI規制のアプローチ (4) その他

	ハードロー	ソフトロー	備考
EU	○（包括的）	○補足的ガイダンス	
米国	△（分野別）	○（包括的）	
カナダ	○（包括的）	○	※AIDA（但し、法案レベル）
ブラジル	○（包括的）	○	※但し、法案レベル
イギリス	○（分野別）	○（包括的）	
イスラエル	○（既存法、分野別）	分野別を補完するガイダンス	
中国	○（生成AI、種類別）	○（包括的）	
シンガポール		○	AI Verifyというツールも
韓国	○（ハイインパクトAI）	○	※但し、法案レベル
日本	○（分野別）	○（包括的）	

*水平・包括的にハードローのアプローチ

cf. A Comparative Framework for AI Regulatory policy, Centre d'Expertise International de Montréal en Intelligence Artificielle(CEIMIA)
https://ceimia.org/en/projet/a-comparative-framework-for-ai-regulatory-policy/

2024年9月末現在

16

スライド 16

2. 各国のAI規制のアプローチ (5) 考察

＜日本＞
- 日本でソフトローが機能する理由
- 日本は失敗をおそれる文化のため、いきなりハードローにすると萎縮効果でイノベーションができなくなるおそれがある。
- もっとも、今後、日本国内のAI事業者が遵守しない、リスクが顕在化すればハードローで対応すべき部分も出てくる。

＜世界的に＞
- どれがいいというわけではない。
- 規制の在り方を考える際に例えば、以下の要因を考慮に入れる必要がある。既存の法律で手当てできているか、社会の特徴（社会的制裁の強さ）、経済的社会的要因、企業文化、技術発達の程度と社会受容等。
  また、学際的 ”Interdisciplinary” な視点が必要。
- 軍事におけるAIの利活用の規制（EUのAI法でも例外）

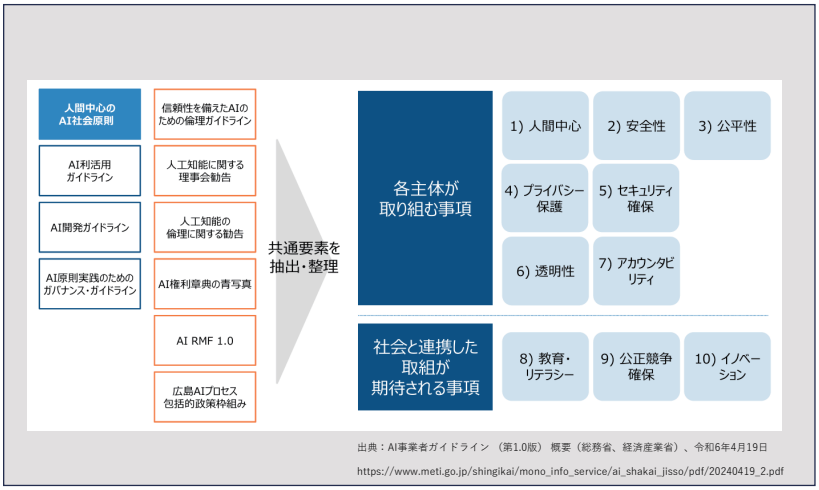
17

スライド 17

3. 相互運用性

• 共通事項

Human-centric
Safety
Security
Fairness
Privacy
Transparency
Accountability



出典：AI事業者ガイドライン （第1.0版） 概要（総務省、経済産業省）、令和6年4月19日
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240419_2.pdf

18

スライド 18

3. 相互運用性

- 国際機関や各国のAI規制は、互いに影響を及ぼしている。

例：G7、G20、 OECD・GPAI、国連、AIセーフティサミット、
AI Safety Summit、広島AIプロセス・フレンズグループ等

19

スライド 19

4. 企業に求められること

＜企業＞
- 日本のAI事業者は、まず、「AI事業者ガイドライン1.0」に取り組む（既に旧ガイドラインから取り組んでいるところが多い）。
  ESG・CSRの観点からも取り組むことが求められている。
- EUで事業展開するところはEUのAI法も見なければならない。
- EUや米国等世界の動向を見ながら開発をする。
  例：EU、米国のNIST、OECD・GPAI

（社内）
- AIポリシーや行動基準の策定⇒リスクの洗い出し、フローの確立、柔軟な体制
  AIガバナンス体制の構築：経営層、リスクマネジメント、法務部、開発現場＋社外委員を含めた「AI倫理委員会」の設置
- AIのデザインから多角的な考察が必要（ダイバーシティ）。
- 販売後も可能な限り情報収集、改善の努力、仕組みが必要。

20

スライド 20

5. 国に求められること

●企業に対する情報提供（ツール、ベストプラクティス）
●分野横断的なAIを所管する組織の必要性
　例：NISCの機能の拡大
●インシデントの共有体制
●政府調達（Procurement）を含む政府のAI利活用基準

21

スライド 21

6. 汎用型AIの規制

●**「弱いAI」から「強いAI」へ**
追跡性(traceability)、説明可能性(explainability)は困難に。制御可能性(Controllability)をどう担保するか？

●AI Safety Institute
英国、米国、日本、カナダ等、続々とAI セーフティ・インスティテュートが作られている。

22

スライド 22

ムーンショット目標１

ムーンショット目標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（PD：萩田 紀博（大阪芸術大学 芸術学部アートサイエンス学科長・教授））

ターゲット

誰もが多様な社会活動に参画できるサイバネティック・アバター 基盤

- 2050年までに、複数の人が遠隔操作する多数のアバターとロボットを組み合わせることによって、大規模で複雑なタスクを実行するための技術を開発し、その運用等に必要な基盤を構築する。
- 2030年までに、1つのタスクに対して、1人で10体以上のアバターを、アバター1体の場合と同等の速度、精度で操作できる技術を開発し、その運用等に必要な基盤を構築する。

注：サイバネティック・アバターは、身代わりとしてのロボットや3D映像等を示すアバターに加えて、人の身体的能力、認知能力及び知覚能力を拡張するICT技術やロボット技術を含む概念。Society 5.0時代のサイバー・フィジカル空間で自由自在に活躍するものを目指している。

サイバネティック・アバター生活

- 2050年までに、望む人は誰でも身体的能力、認知能力及び知覚能力をトップレベルまで拡張できる技術を開発し、社会通念を踏まえた新しい生活様式を普及させる。
- 2030年までに、望む人は誰でも特定のタスクに対して、身体的能力、認知能力及び知覚能力を強化できる技術を開発し、社会通念を踏まえた新しい生活様式を提案する。

https://www8.cao.go.jp/cstp/moonshot/sub1.html

23

スライド 23

ムーンショット目標１

現在、7つのプロジェクトが進行中。

ターゲット1

CA安全・安心確保基盤の構築

ターゲット2

E³LSI課題・政策展開

https://www8.cao.go.jp/cstp/moonshot/gaiyo/ms1_shimpo.pdf

研究開発プロジェクト	PM
誰もが自在に活躍できるアバター共生社会の実現	石黒 浩 大阪大学
身体的能力と知覚能力の拡張による身体の制約からの解放	金井 良太 株式会社国際電気通信基礎技術研究所
身体的共創を生み出すサイバネティック・アバター技術と社会基盤の開発	南澤 孝太 慶應義塾大学
生体内サイバネティック・アバターによる時空間体内環境情報の構造化	新井 史人 東京大学
アバターを安全かつ信頼して利用できる社会の実現	新保 史生 慶應義塾大学
サイバネティック・アバターのインタラクティブな遠隔操作を持続させる信頼性確保基盤	松村 武 情報通信研究機構
細胞内サイバネティック・アバターの遠隔制御によって見守られる社会の実現	山西 陽子 九州大学

https://www8.cao.go.jp/cstp/moonshot/sub1.html

24

スライド 24

ムーンショット目標１

●サイバネティックアバター（CA）とは？

ロボットや3D画像を代理として使用する遠隔アバターだけでなく、ICT（情報通信技術）やロボティクスを活用して人間の身体的・認知的能力を拡張するという概念を含む。

日本語・図による説明：https://www.jst.go.jp/moonshot/program/goal1/files/goal1_explanation1.pdf
英語の定義：https://www.jst.go.jp/pr/announce/20221021/index_e.html

**バーチャル、
メタバース空間における３D画像、
CGなど。**

25

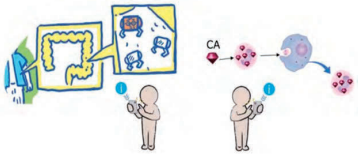
スライド 25

ムーンショット目標１

● サイバネティック・アバター

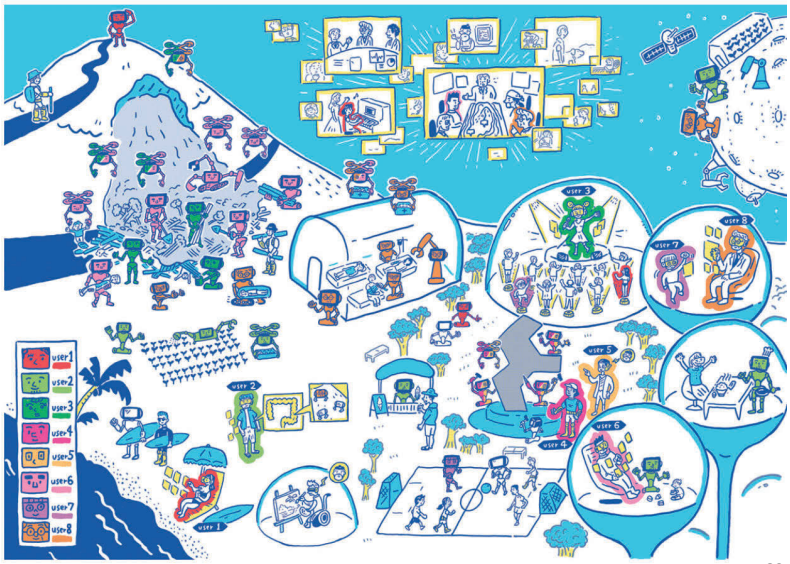
ソシオ CA

体内 CA
細胞内CA



https://www8.cao.go.jp/cstp/gaiyo/yusikisha/20231130/siryo1-1.pdf

2050年はこんな社会になっているかも？



https://www.jst.go.jp/moonshot/program/goal1/illust.html

26

スライド 26

ムーンショット目標１

●E³LSI（read: e-cube-L-si）

ELSI (Ethical, Legal and Social Issues)+Economic（経済的）+ Environmental（環境的）

Ethical, **E**conomic, **E**nvironmental, **L**egal and **S**ocial **I**ssues

ムーンショット目標1「アバターを安全かつ信頼して利用できる社会の実現」プロジェクト（PM：慶應義塾大学 新保史生教授）で造られた用語。

	課題
Ethical（倫理的）	倫理的なCAを開発・利活用するためにはどうすればよいか。
Economic（経済的）	経済への影響は？
Environmental（環境的）	環境への影響は？環境にやさしいエコなCAをどう構築できるか。
Legal（法的）	どのような法規制が必要か？CAに対して既存の法律で足りるか、新法が必要か？ 憲法、競争法、労働法、刑法、個人情報保護、知的財産法等の観点。
Social（社会的）	社会に受容されるためには何が必要か。 社会に対する影響、ヒューマニティ・人類に対する影響。

27

スライド 27

ありがとうございました！

＜お勧めの参考資料＞各国のAI規制について：

A Comparative Framework for AI Regulatory policy: Phase 2 (2024), CEIMIA
-Brazil, Israel, Japan, Singapore and South Korea
*吉永はSteering Committeeの一人として監修。

A Comparative Framework for AI Regulatory policy (2023), Centre d'Expertise International de Montréal en Intelligence Artificielle(CEIMIA)
-Canada, China, EU, UK, USA

https://ceimia.org/en/projet/a-comparative-framework-for-ai-regulatory-policy/

28

スライド 28

＜関連著書＞

- 古川直裕・吉永京子「責任あるAIとルール」
  （一般社団法人 金融財政事情研究会、2024年5月） *Kindleも。
  責任あるAI（生成AI含む）とそれを実現するためのルール、人・企業・国が
  AIとどう向き合っていくべきかを考える手がかりとなるような一般の方向け書籍。
- 商事法務NBL 1269(2024.7.1)号～1278(2024.11.15)号まで9回に渡りEU AI法連載中。
  （吉永は第6回(10/1)、8回(11/1)、9回(11/15)を担当。）
- 「EUのAI法と新興技術規制への視点」【特集：科学技術と社会的課題】、三田評論、2024年8月5日
  https://www.mita-hyoron.keio.ac.jp/features/2024/08-4_2.html
- 「教育現場における生成AIの活用―米国ロースクールにおける生成AIの取り組みの紹介と法学教育における生成AI利活用に関する一考察」【特集／生成AIの法的課題と実務】、有斐閣ONLINE (2024年1月29日）：
  https://yuhikaku.com/articles/-/18714
- 「第2章第３節米国におけるプロファイリング関連制度」福岡真之介・杉浦健二・古川直裕・木村菜生子編著『AIプロファイリングの法律問題-AI時代の個人情報・プライバシー』（商事法務、 2023年10月）
- 『3X-革新的なテクノロジーとコミュニティがもたらす未来』（共著、ダイヤモンド社、2021年）
- 『フロネシス22号- 13番目の人類』（共著、ダイヤモンド社、2020年4月）
- 『フロネシス14号-働きかたの未来』（共著、ダイヤモンド社、2015年12月）
- Jason D. Schloetzer, Kyoko Yoshinaga, Algorithmic Hiring Systems: Implications and Recommendations for Organisations and Policymakers, YSEC Yearbook of Socio-Economic Constitutions 2023- Law and the Governance of Artificial Intelligence, Springer, 2024: https://link.springer.com/chapter/10.1007/16495_2023_61

29

スライド 29

（注 1）令和 7（2025）年 1 月 31 日時点の URL は次のとおり。
https://2021-2025.state.gov/risk-management-profile-for-ai-and-human-rights/

（注 2）令和 7（2025）年 1 月 31 日時点の URL は次のとおり。
https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

報告（3）AI 技術と社会
―倫理・法・社会的課題（ELSI）と諸外国の AI 規制の動向―

慶應義塾大学大学院政策・メディア研究科特任准教授
吉永　京子

最初に、自己紹介をさせていただきます（**スライド 2**）。前職はシンクタンクの研究員で、主に情報通信、メディア、情報セキュリティの分野で、官公庁の委託調査研究を通じて政策立案や法制度改正の支援をしておりました。最後の数年間は、企業向け AI 開発の現場でコンプライアンスやリスクマネジメントを担当し、会社の AI 事業推進ポリシーの策定などにも関わりました。今は幾つかの民間企業において AI 倫理委員会の社外委員も務めております。また、シンクタンク在職時からアメリカの大学で在外研究をしており、そのときに得たネットワークや、昨年から専門委員になりました OECD の GPAI[1]を通じたネットワークを活用して海外の動向も追っているところです。そのような経験から、国の政策形成の観点と企業の AI 開発の観点、また比較法政策の観点から AI ガバナンスを見るようにしております。

本日は、まず AI 規制の簡単な歴史を振り返り、その後各国の AI 規制のアプローチについて特に EU、アメリカ、日本を見てみたいと思います（**スライド 3**）。さらに、相互運用性、企業や国に求められることについてお話しし、最後に、今関わっている科学技術振興機構のムーンショットプログラムの紹介をしながら汎用型 AI の規制について触れたいと思っております。

AI に関してルールが求められるようになった背景には、2000 年代から始まった「第 3 次 AI ブーム」があります（**スライド 4**）。ビッグデータと呼ばれる大量のデータを用いた機械学習が実用化され、その手法の一つとして AI 自ら特徴量[2]を習得するディープラーニング（深層学習）が登場しました。既に、チェスや将棋では AI がプロの人間相手に勝つということが起きており、2016 年 3 月にはグーグル・ディープマインド（Google DeepMind）社が開発した AI の囲碁プログラムが、人間のチャンピオンに勝つという衝撃的なニュースもありました。これは専門家が想定していたよりも 10 年早かったと言われていますが、その頃から様々なリスクが認識されるようになりました。例えば、個人に対する影響だけでなく社会への影響、差別・バイアスの増幅、プライバシーの侵害、AI が仕事を奪うという問題、また AI が出した答えがどのようにして生成されたか人間には理解できないというブラックボックスの問題も指摘されるようになりました。

さらに、2022 年 11 月 30 日（日本時間で 12 月 1 日）に ChatGPT がオープン AI（OpenAI）社からリリースされると、生成 AI がもたらすリスクも問題になってきました。誤情報、ディープフェイク、偽情報を含む違法・有害コンテンツ、バイアス（これは生成 AI 自体が偏見のあるテキストを出すという意味でのバイアスです。）そして著作権、プライバシー、その他諸々倫理的な問題、働くことや教育にもたらす影響も指摘されるようになりました。

例えば、アメリカ国立標準技術研究所（National Institute of Standards and Technology: NIST）

⑴　GPAI（Global Partnership on Artificial Intelligence）は、人間中心の考え方に立ち、「責任ある AI」の開発・利用を実現するため設立された国際的な官民連携組織。
⑵　データや対象物の特徴・特性を定量的に表した数値。

が 2024 年 7 月に公表した「NIST AI 600-1」[3]は、生成 AI 特有の 12 のリスクと開発者が講じる 400 以上の対応策をリスト化した文書です（**スライド 5**）。

リスクに対応する形で、ELSI（Ethical, Legal and Social Issues: 倫理的・法的・社会的課題）が出てきました（**スライド 6**）。ちなみに、Issues の代わりに Implications と書かれている文献も幾つかあります。Implications の場合、「意味合い」という意味になります。

例えば、倫理的課題だとバイアスや公平性、透明性の問題、プライバシー、雇用への影響、また兵器に AI を活用してよいかどうかという問題も出てきます。

法的課題としては、開発者はどこまで責任を負うかという責任分界点が問題です。個人情報保護・プライバシー、知的財産権、また法律で差別が禁止されている場合は差別、いろいろな法律を遵守するという意味での課題も出てきました。

社会的課題としては、人間の職の喪失と労働力の変革、教育への影響、誤情報の拡散によって民主主義自体が揺らぐのではないかといった問題、不平等・格差、社会的孤立、デジタル・デバイド、文化への影響、そして多様性の喪失、更にはヒューマニティへの影響、究極的には人類滅亡の話も出てきました。

冒頭で小塚先生からもありましたとおり、あまり知られていないことですが、日本は世界に先駆けて AI の研究開発に関する原則を提唱し、AI に関する国際ルール作りに貢献しています。2016 年には G7 の議長国として AI の研究開発原則（AI R&D 規則）を提唱しました（**スライド 7**）。その原則に「連携の原則」を追加することで、2017 年には総務省から「AI 開発ガイドライン」が出され、その後 2019 年には「AI 利活用ガイドライン」が出されました。経済産業省も 2022 年に「AI 原則実践のためのガバナンス・ガイドライン（Ver.1.1）」を出しています。

このほか、EU が「信頼性を備えた AI のための倫理ガイドライン」を出したり、OECD が「AI Principles」（人工知能に関する理事会勧告）を出したり、国連もいろいろ出しています。このようにタイムラインで見ると、日本は世界に先駆けて議論してきたわけです。

それでは各国の AI 規制のアプローチとして、まず EU についてお話しします（**スライド 8**）。皆様御存じのとおり、EU では世界で初めて包括的に AI を規制する法律[4]が 2024 年 5 月 21 日に成立し、7 月 12 日に官報に掲載、その 20 日後の 8 月 1 日から順次施行することになりました。EU の加盟国 27 か国で統一的なルールを制定することで EU 域内の市場機能を向上させることが目的としてあります。また、人間中心で信頼できる AI の促進、更に法的確実性をもたらすことでイノベーションを促進したいという思いが EU にはあります。

当初、2021 年に欧州委員会が出した草案は、リスクベース・アプローチというリスクの大小に応じて規制する案になっていました。**スライド 8** のピラミッドを御覧になった方も多いと思いますが、ピラミッドの上から順に、許容できないリスクは禁止する、ハイリスクに関してはいろいろな要求事項や義務を課す、そして限定的なリスクには透明性の原則といったものを課すと書かれていました。最小限のリスクを規制対象外にすることは実際の法律には書かれ

---

(3)　National Institute of Standards and Technology,“NIST AI 600-1 Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile,”2024.7. <https://doi.org/10.6028/NIST.AI.600-1>

(4)　Regulation（EU）2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations（EC）No 300/2008,（EU）No 167/2013,（EU）No 168/2013,（EU）2018/858,（EU）2018/1139 and（EU）2019/2144 and Directives 2014/90/EU,（EU）2016/797 and（EU）2020/1828 (Artificial Intelligence Act)[2024]OJ L2024/1689. <http://data.europa.eu/eli/reg/2024/1689/oj> 同法は、5 月 21 日に EU 理事会で承認され、6 月 13 日に欧州議会と EU 理事会の議長によって署名された。

ていませんが、このような概念として始まりました。

しかし、ChatGPT を始め様々な生成 AI のサービスが出てきたことに伴って汎用目的 AI についての規定をどうしても入れる必要が生じ、一つの章を設けた結果、リスクの大小に応じたアプローチでもなくなりました。現在成立した EU の AI 法を見ると（**スライド 9**）、規制対象として、禁止された AI 行為、ハイリスク AI システム、特定の AI システム、そして汎用目的 AI モデル（いわゆる生成 AI やファンデーションモデルを想定したもの）となっています。

禁止された AI 行為としては、人の潜在意識に働きかけるようなサブリミナルな技法や操作的な技法を用いて人間の行動や意思決定に悪影響を及ぼすこと、ソーシャルスコアリングやプロファイリングをすることなどが書かれており、欧州議会が当初提案したものと比べて、最終的なものには様々な例外事項が付けられています。

EU の AI 法の大部分（8 割ほど）はハイリスク AI システムに関するものです。ハイリスク AI システムに求められる要求事項や、プロバイダー・デプロイヤーなどの義務が記されています。また特定の AI システムに関して、例えば、エンドユーザーがチャットボットやディープフェイクなどの AI に接している場合には、それが AI であることを通知するといった、透明性の要求事項が課されました。汎用目的 AI モデルには、プロバイダーの義務等が書かれ、システミックリスクがあるプロバイダーに関しては、更に重い義務が課されている内容となっています。

本日は時間の関係で詳細は割愛しますが、雑誌『NBL』で現在、何名かの先生方と EU の AI 法を解説する記事を連載していますので[5]、そちらを御覧いただければ幸いです。

続いてアメリカの例です。アメリカは、包括的にはソフトローアプローチ、分野別にはハードローです（**スライド 10**）。州や自治体レベルではハードローもあります。連邦レベルでは 2022 年 10 月にホワイトハウスの科学技術政策局（Office of Science and Technology Policy: OSTP）が、AI 時代におけるアメリカ市民の権利を保護するために、自動化されたシステムの設計、使用、導入に関するガイドとなる「AI 権利章典の構想」[6]を策定しました。その後、バイデン政権が AI 開発先行企業に自主的なコミットメントとして約束させ、2023 年 1 月には NIST が「AI リスクマネジメントフレームワーク 1.0」[7]を公開しました。そして、2023 年 10 月にはバイデン政権が「人工知能（AI）の安心、安全で信頼できる開発と利用に関する大統領令」[8]を出しました。

この大統領令によってアメリカはハードローに舵（かじ）を切ったと誤解されがちですが、そもそも大統領令というものは、ハードローではなく、大統領が各省庁にこういったことについて対応せよという単なる命令にすぎません。各省庁はそれに対してソフトローなガイダンスを作ってもよいということなので、直接 AI を規制するハードローではないということをお話

⑸ 『NBL』1269 号（2024.7.1）から 1278 号（2024.11.15）まで全 9 回にわたり連載された「EU AI 法概説」を指す。このうち吉永氏は第 6 回（1275 号, 2024.10.1, pp.67-78）、第 8 回（1277 号, 2024.11.1, pp.71-79）、第 9 回（1278 号, 2024.11.15, pp.55-63）を担当。

⑹ The White House, "Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People," 2022.10. <https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>

⑺ National Institute of Standards and Technology "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," 2023.1. <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>

⑻ Executive Order 14110 of October 30, 2023: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, *Federal Register,* vol.86 no.210, 2023.11.1, pp.75191-75226. <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence>

ししておきたいと思います。

自治体レベルでは、例えば、ニューヨーク市が採用活動においてAIを利用する場合、少なくとも10営業日前に求職者や被雇用者にAIを利用して採用活動を行うことを通知しなければならないという法律[9]が2023年7月から施行されています。しかし、2024年6月にアメリカに行った際に聞いたところ、あまりうまく機能していないということでした。

そのほか、アメリカ国務省がいろいろなガイダンスを出しており（**スライド11**）、NISTも前述のように生成AIに関するガイドや、アメリカのリスクマネジメントフレームワークと日本の「AI事業者ガイドライン」（後述）の対照表である「Crosswalk2」を公開しています。

州レベルでは、コロラド州がAIに対する包括的な法を全米で初めて制定しました[10]。そのほか、例えば、ユタ州は消費者保護の観点から生成AIの透明性とアカウンタビリティ（人が接しているものがAIだということを通知する等）についての法律を[11]、イリノイ州は採用活動（ビデオインタビュー）におけるAI使用の通知義務に関するAI規制法を策定しました[12]。

ここでお話ししたいことは、最も多くのAI法ができたカリフォルニア州についてです。2023年から2024年にかけての議会において、カリフォルニア州のニューサム知事は17件のAI法案に署名し、17件の法律が制定されました[13]。ディープフェイクやAIの電子透かし、デジタルレプリカなどの透明性を課す法律などがあります。ここで最も論争が多かったSB1047と言われるフロンティアAIに対するAI規制の法案は、知事による拒否権が発動されて成立しませんでした[14]。

最後に紹介する日本は、包括的にはソフトローのアプローチを採っています（**スライド12**）。私も経済産業省「AI事業者ガイドライン検討会」委員として策定に関与しましたが、そこには少子高齢化に伴う労働力の低下でAIを活用する必要があるという背景があります。法律の整備・施行と技術の発達の間でタイムラグが常に生じることや行き過ぎた規制はイノベーションを阻害することから、ソフトローアプローチを採りました。これは「Society 5.0」[15]とも連動しているものです。

加えてあまり知られていないことは、分野別には着々とハードローで既存の法を改正しているということで、既にいろいろな手当てがなされています。先ほど小塚先生からもお話がありましたように、既存の三つのガイドラインを統合して、「AI事業者ガイドライン」が策定され、

⑼　A Local Law to amend the administrative code of the city of New York, in relation to automated employment decision tools (Local Law 144 of 2021). <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9>

⑽　Consumer Protections for Artificial Intelligence. <https://leg.colorado.gov/sites/default/files/documents/2024A/bills/sl/2024a_sl_198.pdf>

⑾　Utah Artificial Intelligence Policy Act (UAIPA). <https://le.utah.gov/~2024/bills/static/SB0149.html> なお、法律に統合されたのは次の箇所である。Chapter 72 of Title 13 in the Utah Code. <https://le.utah.gov/xcode/Title13/Chapter72/C13-72_2024050120240501.pdf>

⑿　イリノイ州、メリーランド州における法律の例として以下。Artificial Intelligence Video Interview Act. <https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=4015&ChapterID=68>; Labor and Employment - Use of Facial Recognition Services - Prohibition. <https://mgaleg.maryland.gov/mgawebsite/Legislation/Details/HB1202?ys=2020RS>

⒀　成立した法律とその概要の一部については以下の二つの記事にまとまっている。北村弥生「【アメリカ】選挙広告におけるAI等による実質的に虚偽のコンテンツの拡散を規制する法律の制定（カリフォルニア州）」『外国の立法』No.301-2, 2024.11, pp.6-7. <https://doi.org/10.11501/13783827>; 北村弥生「【アメリカ】カリフォルニア州におけるディープフェイクを規制する法律」『外国の立法』No.302-1, 2025.1, pp.2-3. <https://doi.org/10.11501/13979498>

⒁　Safe and Secure Innovation for Frontier Artificial Intelligence Models Act. <https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202320240SB1047>

⒂　「Society 5.0」内閣府ホームページ <https://www8.cao.go.jp/cstp/society5_0/>

4 月 19 日に公表されました（**スライド 13**）。主な特徴は（**スライド 14**）、リスクベース・アプローチ、アジャイル・ガバナンス、事業者の自主的な取組の支援、そして国際的な議論との協調として、EU の AI 法なども入れ込みながらコンセプトが作られている点です。またこのガイドラインは、マルチステークホルダーの観点から様々な方と議論して策定されており、なおかつ Living Document として毎年改正する方向です。加えて、AI セーフティ・インスティテュート（AISI）が積極的に様々なガイダンスを出しています（**スライド 15**）。

海外の状況を見ると（**スライド 16**）、EU は包括的な AI 規制を策定しました。カナダ・ブラジル・韓国もその方向に向かっていますがまだ法案レベルで、かつ日本やアメリカがソフトローということから少しトーンダウンして進んでいると聞きました。**スライド 16** で黒字の部分は、分野別で規制をしようとしているところです。

国際学会でよく言っていることですが、日本でソフトローが機能する理由は、日本にはお上に言われたら従うという法文化があることです（**スライド 17**）。そのため、諸外国に比べて日本は法的拘束力がなくても従います。そこが良いところだと思っています。また、日本は社会的制裁が強い国で、例えば、個人情報が少し漏洩（ろうえい）しただけでも、メディア等を通じて社会から強い非難を受けるため、企業はその信頼や評価を落とさないように、CSR（企業の社会的責任）の観点から対応する必要があるというインセンティブがあります。また日本の事業者には失敗を恐れる文化があるため、いきなりハードローにすると萎縮（いしゅく）効果でイノベーションできなくなる恐れもあると思っております。ソフトローからスタートして、徐々にリスクが顕在化すればハードローで対応するのが良いと私は思っています。

世界的に見てどれが良いというわけではなく、やはり国のいろいろな事情、つまり社会的経済的要因、企業文化、技術の発達程度に応じてソフトローとハードローのどちらを採用するか選択すればよいと思っています。また、学際的な視点も必要です。なお、軍事における AI の利活用の規制については、今のところ EU の AI 法でも例外になっています。

相互運用性については、どの国も、**スライド 18** に掲げたようなプリンシプルを包含しています。また国際機関や各国の AI 規制は、**スライド 19** に挙げているような団体を通じて、互いに影響を及ぼしています。

最後に、日本の企業と政府に求められることについて触れます。日本の企業はできるところからやっていき、EU で事業展開をする場合は EU の AI 法も見ることが求められます（**スライド 20**）。また、国については、企業に対する情報提供もそうですが、分野横断的な AI を所管する組織の必要性があると思います（**スライド 21**）。例えば、内閣サイバーセキュリティセンター（NISC）は省庁横断的なセンターですが、その機能を拡大する手もあるのではないかと思っています。このようなインシデントの共有体制や政府調達の基準も今後日本では必要になってきます。

汎用型 AI の話題（**スライド 22** 以降）はパネルディスカッションのときに触れたいと思います。

（よしなが　きょうこ）