

# 人工知能 (AI) の利用と規制

## —よりよい社会を目指して—日本の場合

岩本 誠吾

### The Uses and Regulations of Artificial Intelligence (AI)

#### —Aiming for a Better Society— in the Case of Japan

IWAMOTO Seigo

1. 問題提起——日本の現状
2. 日本の AI 規制動向
3. AI 規制枠組みの事例研究：EU の場合
4. 2023 年度での日本の AI 規制取組
5. 日本による AI 規制の展望

#### 1. 問題提起——日本の現状

日本は、近未来社会を Society 5.0<sup>1)</sup> と位置付けて、サイバー空間（仮想空間）とフィジカル空間（現実空間）を高度に融合させたシステムにより、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」を建設しようと計画しています。Society 5.0 では、サイバー空間でビッグデータが人工知能 (AI) により解析され、その結果がロボットなどを通じて人間にフィードバックされます。AI は、Society 5.0 実現には必要不可欠な存在であると言えます。

その AI は、特化型 AI (specialized AI)<sup>2)</sup> ではありますが、既に身近な領域で活用されています。たとえば、お掃除ロボット、医療分野での癌の早期発見、見守りや話し相手用の高齢者用介護ロボット、ビッグデータ解析による防災情報の提供などです。最近では、2022 年 11 月 30 日にアメリカのオープン AI 社が生成 AI のチャット GPT を公表しました。直ぐに日本でも多くの人がそれを自動翻訳や文書作成の事務処理などに活用しています。他方で、その成果物が本当に正しいのか、著作権を侵害していないのか、より具体的なデメリットやリスクが認識されるようになりました。

未来学者のレイ・カーツワイル<sup>3)</sup> によれば、2045 年には AI は人間の知能を凌駕すると予言してい

ます。今後、AIの研究開発が進めば、汎用型AI（Artificial General Intelligence, AGI）<sup>4)</sup>の登場も予想され、我々は、AIにより大きな期待とともに深刻な不安を抱くようになりました。AIの科学技術があまりにも急激な進展を遂げているのに対して、現代の法制度はそれに全く対応しておらず、現代社会はAIの無法状態にあると言えます。現時点では、AIの発展をそのまま放任するかまたは禁止するかという段階の議論は過ぎて、我々人間とAIが如何に共存・共生するのかという段階にあると思います。すなわち、AIを使用する場合には、必然的に、光と影、メリットとデメリットがあり、その現状にどう対応すべきか、これが現代国際社会の喫緊の課題となっています。

この報告では、日本を事例として、AIはどのような方向に進み、どのような方法で利活用すべきなのか、逆に言えば、どのようにして禁止・規制すべきか、を検討したいと思います。

## 2. 日本のAI規制動向

まず、2017年に、日本の人工知能学会が民間学術団体として「倫理指針」を作成しました。次に、日本政府が2019年3月に「人間中心のAI社会原則」を公表しました。

ここでは、人間の尊厳が尊重される社会を基本理念として、AI社会原則として、「人間中心の原則」、「セキュリティ確保の原則」、「公平性、説明責任及び透明性の原則」などが規定されています。同年5月に、日本も加盟国である経済協力開発機構（OECD）が「人工知能に関する理事会勧告（AI勧告）」により「信頼できるAIの責任ある運営原則」（OECD原則）<sup>5)</sup>を採択しました。

そこには、5つの原則が含まれています。①AIは、包摂的成長と持続可能な発展を促進し、人間や地球環境に利益をもたらすものでなければならない。②AIは、法、人権、民主主義、多様性を尊重し、公平公正で、偏りや不平等を最小限に抑えて、必要に応じて人的介入ができるようにすべきである。③AIの意思決定や行動は、透明性を確保し、責任ある情報開示を行うべきである。④AIの設計者と運用者は責任を持ち、リスク管理と安定・安全性を確保しなければならない。⑤AIの開発、普及、運用に携わる組織および個人は、上記の原則に則ってその正常化に責任を負うべきである。

特に、AIシステムは、②で人間による最終的な意思決定の余地を残し、かつセーフガード（安全装置）を実装すべきであると明記しています。

このOECDの「人間中心アプローチ（human centric approach）」や5つの原則を含む「信頼できるAI（trustworthy AI）」という概念は、2019年の欧州連合（EU）の「信頼できるAI倫理指針」でも含まれ、かつ、同年のG20大阪首脳宣言でも追認されています。すなわち、OECD、G20、EUの50以上の国家は、それらの概念をAI倫理規範として是認しています。

### 3. AI 規制枠組みの事例研究：EU の場合

AI は、規制内容に係る 5 原則とは別に、規制枠組みが別次元の課題として存在します。欧州連合 (EU) は、2021 年に世界初の AI 利用に関する規制枠組みの「AI 規制法案」を提案しました。ここでは、AI 規制枠組みの在り方の参照事例として EU の「AI 規制法案」を取り上げます。それによれば<sup>6)</sup>、リスクを基準に AI システムを 4 段階に区分し、それぞれの段階での対応の仕方を決めています。これは、リスク・ベース・アプローチ (risk-based approach) と呼ばれています。

まず、容認できないリスク (unacceptable risk) とは、人間の基本的権利を侵害するような場合を指し、その利用は禁止される。第 2 に、高リスク (high risk) とは、市民の生活や健康を脅かす恐れがある、また安全性に問題が懸念される場合を指し、その利用には厳格な要件が課される。第 3 に、限定的リスク (limited risk) とは、深刻な危険性はない場合を指すが、その利用には、透明性に関する特定の要件を満たす必要がある。第 4 に、最小のリスク (minimal risk) とは、リスクがごく僅かか、リスクを伴わない場合には、そのまま利用する。

このように、AI システムに内在するリスクの程度に応じて、①禁止、②規制、③透明性義務のみ、④規制なし、という対応措置が採られることとなります。EU は、このアプローチを、AI の開発者と利用者に対して AI のイノベーションを促進するとともに、安全性の向上と基本的権利の保護を両立させるものであると位置付けています。

今後、EU は、具体的にどのような AI システムがどの段階に分類されるのか、どのような規制措置が必要なのか、を詳細に検討していく必要がありますが、AI 規制の枠組みとしては、日本にとっても非常に参考になる基準であると考えます。

### 4. 2023 年度での日本の AI 規制取組

今年の 2023 年に、G7 主要国首脳会議が、議長国の日本で開催されました。G7 デジタル相会合は、4 月に「責任ある AI と AI ガバナンスの推進」を謳った閣僚宣言が採択されました。そこでは、「OECD の AI 原則に基づき、人間中心で信頼できる AI を推進し、AI 技術がもたらす全ての人の利益を最大化するために協力を促進するとのコミットメントを再確認」しています。なお、閣僚宣言 (45 項) は、AI の政策と規制は「リスクベース」で指向しなければならないと再確認し、EU のリスク・ベース・アプローチを採用しています。

ただし、G7 の中でも、たとえば、生成 AI に関する厳しい規制を目指す欧州とそうではない国 (たとえば、日本や米国) との立場に若干の違いがあります。閣僚宣言 (43 項) でも、5 月の G7 広島サミット首脳宣言 (38 項) でも、「信頼できる AI という共通目標へのアプローチと政策手段が G7 諸国間で

異なる場合がある」ことを認識して、AI ガバナンスの枠組み間の相互運用性を重要視しています。そして、特に、生成 AI の国際ルール作りの議論のために、「広島 AI プロセス」を年内に創設すると表明しています。

その流れを受けて、5月に日本の AI 戦略会議は、AI の論点整理をしました。そこでは、7つの懸念リスク、①機密情報の漏洩や個人情報の不正利用、②犯罪の巧妙化・容易化、③偽情報による社会の不安定化、④サイバー攻撃の巧妙化、⑤学校現場における生成 AI の扱い、⑥著作権侵害、⑦ AI による失業者の増加を挙げています。その対応策として、過度な規制を避けつつ、AI 開発者、サービス提供者、利用者が自らリスクを評価し、既存の法令やガイドライン<sup>7)</sup>の遵守によるガバナンスを促しています。

## 5. 日本による AI 規制の展望

AI の議論は、世界も日本も、AI 倫理原則はなぜ必要か (why)、AI 倫理原則は何か (what) から、

表【日本の AI 規制動向】 著者作成 (2023 年 9 月 12 日現在)

年	月日	日本、日本参加機構・会議その他一般的な AI 関連事項
2017	2/28	日本の人工知能学会が人工知能学会倫理指針を承認
	7/28	日本の AI 開発ガイドライン
2019	3/29	日本・内閣府「人間中心の AI 社会原則」
	4/8	EU「信頼できる AI 倫理指針」
	5/22	OECD「AI に関する理事会勧告」の OECD 原則
	6/29	G20 大阪首脳宣言 (G20AI 原則)
	8/9	日本の AI 利活用ガイドライン
2020	10/20	欧州議会 AI 倫理枠組み決議
	10/20	G20 リヤド首脳宣言 (G20 AI 原則推進国内政策例)
2021	4/21	EU「AI 規制法案」公表
	11/23	UNESCO「AI 倫理勧告」採択
2022	11/30	オープン AI が AI チャットボットのチャット GPT (生成 AI) 公表
2023	4/30	G7 デジタル・技術相会合閣僚宣言 (AI 利用 5 原則合意)
	5/20	G7 広島首脳コミュニケ (広島 AI プロセス)
	5/26	日本の AI 戦略会議の AI 論点整理
	6/14	欧州議会が AI 規則案採択
	7/21	米政府が AI 企業 7 社と AI の安全性の自主規制に合意 (9/12 に 8 社参加)

AI原則の実現のための規制・ガバナンスの在り方（how）に移行しつつあります<sup>8)</sup>。AIの規制またはガバナンスの形式は、4で述べたように、法規制（ハード・ロー）なのか、法的拘束力のない倫理規範や行動指針（ソフト・ロー）なのか、AI企業やAI利用者の自主規制なのか、それらをどのように組み合わせるのが焦点になっています。

法的義務を課す法規制は、AI規則の実効性を高めるには最適ですが、多数国による規則策定過程に時間がかかり、その実現可能性が低くなります。他方、非法的な行動準則や行動指針の場合は、緩やかな規制ゆえに、それへの賛同国が増える利点がありますが、その指針の履行確保には不透明さが残ります。

そのような状況の中で、日本は、「広島AIプロセス」を推進し、AIの国際ルール作りやG7諸国間の合意形成に貢献するために、まず、厳格な法規制（ハード・ロー）よりも、行動指針や自主規制の組合せ（ソフト・ロー）を選択することになるのではないかと推測します。日本は、広島AIプロセスの推進と同時に、別のAIガバナンス枠組みとの相互運用性の調整を図ることになるかと思っています。

我々は、AIの恩恵を享受すると同時に、発生し得るリスクを回避するために、AIの急激な技術革新に伴走しながら、AI規制・ガバナンスの実現のための努力をする必要があります。その場合に、我々は、あくまで主体は人間であり、AIは補完的役割に徹するとの大原則を再認識しなければいけないことは言うまでもありません。（2023年10月9日）

## 注

- 1) 狩猟社会を Society 1.0、農耕社会を Society 2.0、工業社会を Society 3.0、情報社会を Society 4.0と位置付ける。Society 5.0は、2016年に閣議決定された「第5期科学技術基本計画」において内閣府が提唱した概念である。
- 2) 個別の領域に特化したAIで、限定された範囲の処理を行うシステムを指す。
- 3) Ray Kurzweil, *The Singularity Is Near: When Humans Transcend Biology*, Viking Press, 2005, P. 136. 人間知能と人工知能が逆転する時点を技術的特異点（technical singularity）という。
- 4) 人間のように推論して多様な状況に対応できる人工知能を言う。
- 5) Principles for responsible stewardship of trustworthy AI. 本原則は、2020年3月31日時点で遵守国はOECD加盟国39カ国と非加盟国8カ国の計47カ国である。
- 6) 駐日EU代表部公式ウェブマガジン『EU MAG』2021年8月5日 vol. 83（2021年夏号）参照。
- 7) AIネットワーク社会推進会議は、たとえば、2017年7月28日にAI開発者用に「AI開発ガイドライン案」を、2019年8月9日にAI利用者用に「AI利活用ガイドライン」を作成した。
- 8) 江間有沙『AI社会の歩き方—人工知能とどう付き合うか』DOJIN 選書080、2019年、111頁。