

# AIガバナンスに関する提言

Ver 3.0

GOVERNANCE



デジタル政策フォーラム



## 目次

---

はじめに ..... p.3

### AI ガバナンスに関する提言 Ver 3.0

基本的考え方 ..... p.4

1 リスクの最小化 ..... p.5

- (1) リスク管理のあり方
- (2) 規制のあり方と実効性の確保
- (3) 外的リスクに対する脆弱性対策
- (4) 生成物の取り扱い

2 利便性の最大化 ..... p.13

- (5) AI の積極的活用

3 健全な市場環境の整備 ..... p.15

- (6) 健全なエコシステムの構築 – 競争政策
- (7) 産業振興とグローバル連携 – 産業政策
- (8) 国際的コンセンサスの醸成 – 外交政策
- (9) AI がもたらす広範な影響に関する議論
- (10) 倫理的な問題への対処

今後の作業計画 ..... p.22

### AI ガバナンスを巡る論点 2025

有識者インタビュー ..... p.23

<最前線>企業の AI 利活用 ..... p.83

<特集> AI ×クリエイター ..... p.107

鼎談 デジタル政策フォーラムのアクション ..... p.117

おわりに ..... p.124

# はじめに

デジタル政策フォーラム (DPFJ) は、2024 年春から AI ガバナンスに関する議論を続けてきた。AI ガバナンス、すなわち急速に進化する AI の技術的・社会的なコントローラビリティをどのように確保していくかという課題を主テーマに多数の有識者や政府・企業関係者と議論を深め、2024 年 7 月に第一弾の提言、同年 12 月に第二弾の提言を公表した。

第一弾の提言では、AI ガバナンスとして、(1)AI リスクの最小化、(2)AI の利便性の最大化、(3)健全な市場環境の整備という3項目を議論の柱として設定した上で主要論点を整理した。この枠組みは第二弾、そして今回の第三弾においても踏襲している。

第三弾をまとめる過程で実感したのは、AI の実利用が企業や個人のレベルで急速に進んでおり、これに伴い、AI を巡る論点も具体化している点である。例えば、企業ヒアリングでは、AI に対する経営層の理解が深まり、AI 活用が経営戦略の主要アジェンダになっていること、また AI による経営効率化（コスト削減）だけでなく、新しい事業を創出するという付加価値の向上の視点が数多く見られるようになってきていることが明らかになった。

また、今回の提言の特徴の一つとして、AI を巡る政策議論が、産業政策、競争政策、科学技術政策、安全保障政策、外交政策など幅広い分野に及んでおり、かつそれぞれの政策領域の相互関連性が強まっているという点も挙げられる。これは AI ガバナンスのあり方が国益に直結する重要度の高い政策テーマになっていることを意味している。

さらに、AI の実利用が進展する中、議論すべき論点の対象範囲も大きく拡大している。具体的には、AI を巡る政策議論にとどまることなく、AI は社会経済の何を変えるのか、例えば AI と創造性（クリエイティビティ）の関係、デジタルデモクラシーの可能性、宗教と AI の関係性など、今回の第三弾の提言はこれまでの議論を大幅に拡張したものとなっている。かつてインターネットは時間と距離の制約を越えることを可能にした。今、AI は言葉の壁を越え、組織の壁を越え、サイバーとフィジカルの壁を越えることを次々に可能にしている。データ駆動型社会の加速化装置である AI の実装がどこまで大きなインパクトをもたらすのか。これまでの IT 中心のテックコミュニティだけでなく、幅広い領域の専門家とのコラボレーションも今後 AI ガバナンスを巡る議論において重要になってくるだろう。

AI を巡る議論で注意を要するのは、議論の対象である AI そのものが引き続き急速に進化し続けているということだ。米国の著名な AI 研究者であるエリーザー・ユドコフスキー博士は「AI の最大の危険は、人々がそれを完全に理解していると信じてしまうことだ」と警鐘を鳴らす。DPFJ は AI を巡る技術動向や市場変化、さらに社会経済構造に与える影響などを定点観測的に分析し、第四弾の提言としてまとめる予定である。

最後に、第三弾提言の作成に際しご協力をいただいた有識者や企業等の関係者の皆様（本書にインタビュー記事等を掲載）にこの場を借りて深く感謝申し上げたい。

谷脇 康彦  
デジタル政策フォーラム 代表幹事  
2026 年 3 月

# AI ガバナンスに関する提言 Ver 3.0

2026年3月  
デジタル政策フォーラム

## 基本的考え方

生成 AI を巡る技術開発は著しい速度で進んでおり、生成 AI を社会経済システムに実装する動きも加速化している。こうした中、生成 AI を制御するためのルールづくりは、各国における法的枠組みといった総論の段階から具体的な政策のあり方を巡る議論へと比重を移しつつある。

本文書<sup>1</sup>では、こうした生成 AI を巡る動向を念頭に置きつつ、検討の基本的な視点として、

1. AI のリスクの操作可能性の確保（リスクの最小化）
2. AI の利便性を最大限享受できる環境の整備（利便性の最大化）
3. 上記環境を自律的に実現する市場の創出（健全な市場環境の整備）

という3つの目的を均衡ある形で実現する AI ガバナンスに関する議論を可能な限り広範な視点で整理する。

AI ガバナンス、すなわち AI 技術の制御可能性を継続的に維持するための仕組み<sup>2</sup>については、AI がもたらす便益とリスクのバランスを常に比較しながら議論する必要がある。

AI は社会のあらゆる領域で生産性や創造性の向上に寄与するものであり、パーソナル化（インテリジェンスの分散化）を通じて個人のデータ利用に係るデータ主権（data sovereignty）を技術的に担保しつつ、利便性の高いサービスが享受できるようになるなど、多種多様な便益をもたらす。

他方、人権侵害などの被害を深刻化するリスク、人間による制御可能性が失われるリスク、さらには AI が人間を代替することで生じるリスクなどが考えられる。ちなみに、こうしたリスクについては可能な限り技術的解決を目指し、拙速に規制を導入することはイノベーションを促す観点から適当ではない。

デジタル政策フォーラム（DPFJ）は、2024年7月、AI ガバナンスの枠組みの構築に向けた論点を提言 Ver 1.0 として公表した<sup>3</sup>。さらに、同年12月にはより詳細な論点の整理等を行い、特に AI 法を巡る議論が具体化していたことを踏まえ、法的枠組みのあり方や AI 総合戦略の推進の必要性などについて提言 Ver 2.0 として公表した<sup>4</sup>。

しかし、AI 技術が引き続き急速な進化を遂げ、議論の対象としての AI そのものが変化し続けている状況にある。

そこで、提言 Ver 2.0 以降も続けてきた有識者インタビューや企業関係者からのヒアリング、関連会合

1 本文書は谷脇康彦デジタル政策フォーラム（DPFJ）代表幹事が監修した。

2 総務省・経済産業省「AI 事業者ガイドライン（第 1.1 版）」（2025年3月）総務省・経済産業省において、AI ガバナンスは「AI の利活用によって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ、そこからもたらされる正のインパクト（便益）を最大化することを目的とする、ステークホルダーによる技術的、組織的、及び社会的システムの設計並びに運用」と定義されている。

[https://www.soumu.go.jp/main\\_content/001002576.pdf](https://www.soumu.go.jp/main_content/001002576.pdf)

3 [https://www.digitalpolicyforum.jp/wp-content/uploads/2024/06/240701\\_AI01.pdf](https://www.digitalpolicyforum.jp/wp-content/uploads/2024/06/240701_AI01.pdf)（以下、提言 ver1.0）

4 <https://prtimes.jp/main/html/rd/p/000000011.000131931.html>（以下、提言 ver2.0）

の開催<sup>5</sup>等を踏まえ、かつ最近の海外動向なども反映させつつ、改めてAIガバナンス全般にわたる論点整理を行い、「AIガバナンスに関する提言 Ver 3.0」として整理する。

ここにはAIガバナンスを確保するための法的枠組み、リスク管理手法、AIを巡る競争政策、産業政策、外交政策及び安全保障政策、その他AIがもたらす社会構造変化のインパクト分析を含む多種多様な項目が取り上げられており、かつこれらの項目が有機的に相互作用している点に注目する必要がある。AIガバナンスを論じるには俯瞰的な「鳥の目」が求められる。

なお、本文書では現時点で一般向けに提供されている生成AIを主として念頭に置きつつ議論を進めることとし、汎用人工知能（Artificial General Intelligence）は一部を除き対象としない。

# 1 リスクの最小化

## (1) リスク管理のあり方

### < AIの段階別リスク管理の困難性 >

AI管理の手法としてリスク（人の生命や基本的人権に与える負の影響などを含む）を数段階に分けて管理する手法が存在する。例えば、EUのAI法ではリスクを4段階に分類<sup>6</sup>している。

これはAIモデルが持っているリスクを深刻度に応じて段階別に管理しつつ、これを規制の度合いにリンクさせるものである。しかし、この手法の場合、コントロールすべきリスクの範囲をどう画定するか、またリスクをどのような基準でランク分けするのかといった具体的なリスク管理の手法に加え、リスク判断の主体、当該主体の判断の的確性を第三者に明示する手法（説明責任）等が確立しているとは言い切れない。

この点、将来的にはAIのシステムログ（動作履歴）を記録・解析することでAIのリスクをスコアリングし、これを公表する仕組みが構築されることとなれば、自らが許容できるリスクに応じてAIを選択することができるようになる可能性がある。すなわち、当該AIから得られる便益（ベネフィット）とリスク（コスト）を比較考量し、各利用者が自らの用途に適した（パーソナル）AIを選択できる仕組みが構築できることを念頭に置きつつ、国際機関におけるAIの標準化（リスク評価の手法を含む）等の議論に積極的に参画していくことが必要である。

またAIモデルのリスク評価を規制の強弱とリンクさせるアプローチは、リスクそのものが変化し得る可能性を踏まえれば、規制適用の予見可能性を損なう可能性がある。この点、欧州においては高リスクAI

5 DPFJ主催オープンカンファレンス「デジタル政策の論点2026」（2026年2月）において、AIガバナンスが主要テーマの一つとして取り上げられた。

6 欧州AI法は、AIのリスクを①許容できないリスク（人の生命や基本的人権に対する直接的脅威を及ぼすものとして開発を禁止）、②高リスク（事前の適合性評価、データベースへの登録等の義務）、③限定リスク（AIとのやり取りであることを利用者に知らせる透明性確保の義務）、④最小リスク（規制なし）の4種類に分類している。

の技術要件を緩和する「デジタルオムニバス提案」<sup>7</sup>が行われているところであり、今後の議論の動向を注視する必要がある。

事実、AIの抱えるリスク源は多様であり、その全容を把握することが困難である。例えばMIT調査<sup>8</sup>によればAIを巡り700を超えるリスクが存在するところであり、そのすべてを念頭に置いたリスク管理を制度として実装することには大きな困難を伴う。また、本調査では「開発後（post-deployment）のリスク」がリスク全体の65%を占めるとの指摘があるなど、リスクが時間の経過とともに動的・質的に変化する。

無論、AIのリスク管理そのものは極めて重要であり、国内においても産学官の連携によりAIリスクに関するレポジトリーの作成・分析等を積極的に推進すべきである。

## <主体別のリスク管理>

上記を踏まえつつ、AIのリスク管理については、

- ・ AIの開発者
- ・ AIを実装したサービス提供事業者
- ・ エンドユーザー

という3つの主体に分けて検討することが望ましい。ちなみに、AIのリスクを検討する際、開発段階においてAIが内包する可能性があるリスクと、サービス提供段階においてAIが有することとなるリスク（AIを実装したサービスの提供・利用の方法等によって顕在化する可能性があるリスク。例えば誤情報・偽情報の生成・流布などが含まれる）の2つが考えられるが、特に後者のリスクについては、当該リスクがAIによって初めてもたらされたものであるのか、それとも従来から存在しているリスクがAIによって顕在化・増幅したものであるのか等について慎重に議論する必要がある。

主体別リスク管理のうち、AIの開発者によるリスク管理は、開発時に留意すべき事項を限定的に列挙する“Do Not List”アプローチに止めることとし、今後、AI開発において具体的な問題が発生した場合にはその時点で対処することを基本としつつ、定期的なモニタリングを行うことが考えられる。

具体的には、例えば欧州評議会のAI条約（2024年9月）<sup>9</sup>等を参照しつつ、「AIシステムのライフサイクルにおける活動が“人権・民主主義・法の支配”と十分な整合性を確保すること」などを原則とするこ

7 2025年11月、欧州委員会は「デジタルオムニバス提案」を公表した。欧州域内における規制コスト負担を軽減することを目的とするオムニバス提案は10の領域に及び、デジタル分野もその一つ（オムニバス7）とされている。この提案の中では一定の要件（再識別リスクの適正な管理、データ主体の異議申し立ての機会の確保、データの最小化等）に適合する個人情報AI学習データとして利用することを合理的利益として許容するほか、高リスクAIに関する技術要件の緩和を検討項目として挙げている。特に技術要件の緩和についてはAI法の該当部分の施行を最大で2027年末までに延長（当初予定は2026年8月）することが提案されている。

なお、本提案は現時点では草案段階であり、欧州議会及び欧州理事会との調整を経て立法化がなされるものと見込まれている。

（出典）European Commission “Simpler EU digital rules and new digital wallets to save billions for businesses and boost investment”（November 2025）。

[https://ec.europa.eu/commission/presscorner/detail/en/ip\\_25\\_2718](https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2718)

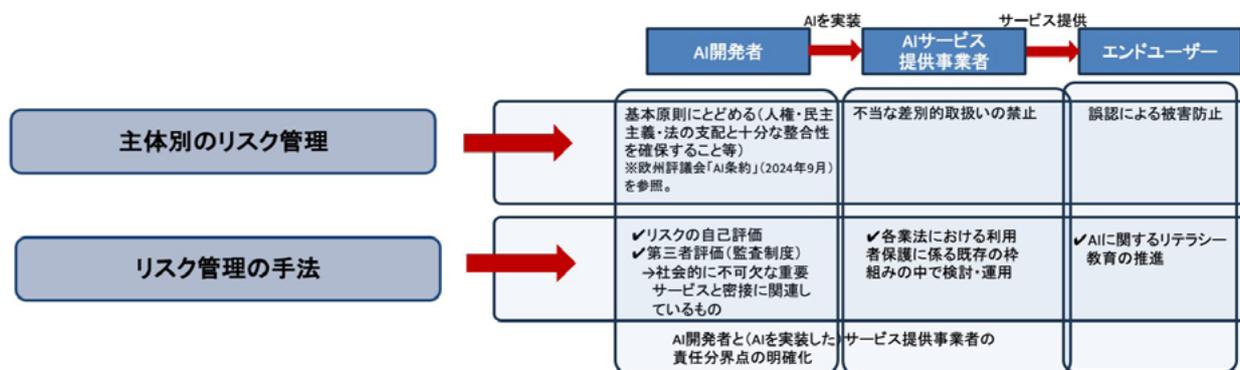
8 P. Slattery et al. “Global AI adoption is outpacing risk understanding, warns MIT CSAIL”（MIT CSAIL News, August 14, 2024）

<https://www.csail.mit.edu/news/global-ai-adoption-outpacing-risk-understanding-warns-mit-csail#:~:text=Global%20AI%20adoption%20is%20outpacing,it%20remains%20current%20and%20relevant.>

9 2024年9月、「AI並びに人権、民主主義及び法の支配に関する欧州評議会国際条約」（Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law）に米国・EUを含む15を越える国・地域が署名（日本は2025年2月に署名（※））。

（※）：[https://www.mofa.go.jp/mofaj/press/release/pressit\\_000001\\_01725.html](https://www.mofa.go.jp/mofaj/press/release/pressit_000001_01725.html)

## AIに関するリスク管理



とが考えられる。

また、AIを実装したサービス提供事業者によるリスク管理についても、可能な限り限定的であることが望ましい。例えば電気通信事業法第6条<sup>10</sup>が規定するように、サービス提供における不当な差別的取り扱いを禁止する等の規律にとどめることが望ましい。

AIに関連して不当な差別的取り扱いを禁止するのは、AIを活用することで個別医療など個人情報を活用したきめ細かいサービスの提供が可能となる一方、個人の特性に応じて個別化(personalization)されたサービス提供ではなく、合理的な根拠に欠ける差別(discrimination)とならないようにすることが人権保護の観点から求められるからである。

なお、AIを組み込んだサービスを利用者に提供する場合、AIの開発者とサービス提供事業者との間の責任分界点についてもサービス提供の前段階において予め明確にしておくことが利用者保護の観点から求められる。

さらに、エンドユーザー(中小企業等を含む)におけるリスク管理については、AIのリスクについて正しく理解するためのリテラシー教育が求められる<sup>11</sup>(項目(5)を参照)。

### <リスク管理の手法>

リスク管理は既述のAIリスクリポジトリの作成・分析の結果等を踏まえて行うべきであるが、その際、リスクの自己評価あるいは第三者評価(例えば監査もしくは認証制度)の適用の可否について検討する必要がある。

10 電気通信事業法第6条は「電気通信事業者は、電気通信役務の提供について、不当な差別的取扱いをしてはならない。」と規定している。

11 例えば、2024年9月、韓国において性的なディープフェイク画像や動画の所持・視聴を処罰する「性暴力犯罪処罰などに対する特例法」の改正法が成立している。また、Freedom Houseは報告書“Freedom on the Net 2025”

(※)において、AI生成コンテンツに関する規制が世界中で急速に導入されてきており、規制のアプローチは国によって大きく異なり、権威主義国家では検閲や監視の強化に力を入れる(国の方針に沿わない生成物の生成を禁止する)一方、民主主義国家では子供の保護と表現の自由の適正なバランスの確保に苦慮していると指摘している。

(※) Vesteinsson, Baker, Brody, Funk, Grothe, Slipowits eds. Freedom on the Net 2025, Freedom House, 2025 www.freedomonthenet.org

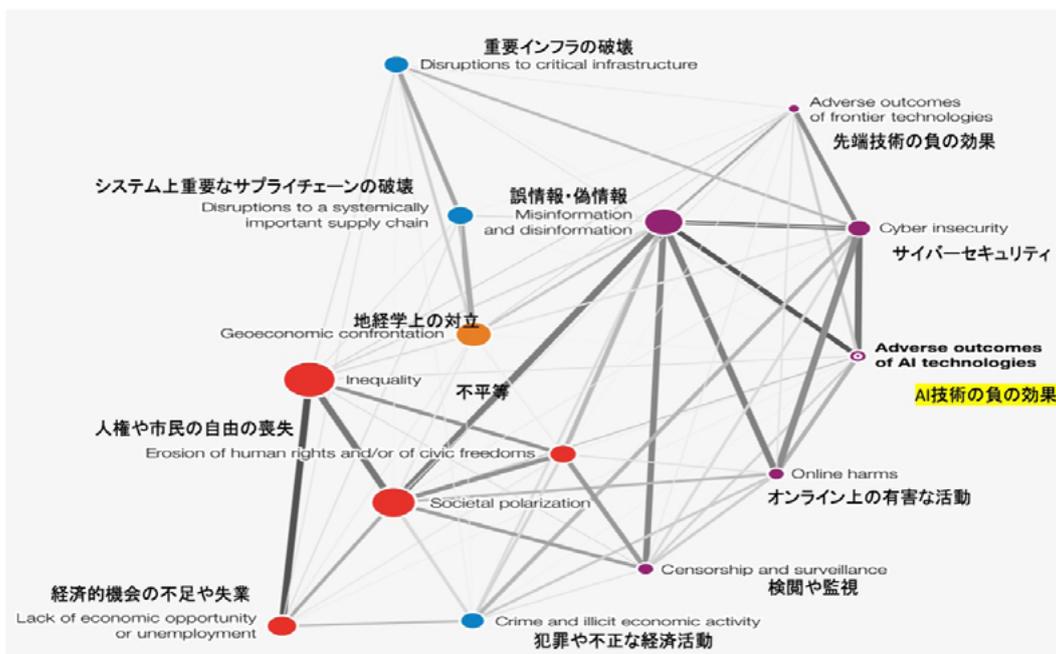
具体的には、AIの開発者においては開発者自らによる自己評価を基本とし、社会的に不可欠な重要サービスと密接に関連しているものについては第三者による監査制度を組み合わせることが考えられる。

またサービス提供事業者においてはAIの機能だけを抜き出して評価することは困難であることから、各業法における利用者保護に係る既存の枠組みの中で検討・運用すべきであり、AIの活用を契機として規制の上乗せを行うことは適切ではない。

## <システミックリスク分析の必要性>

AIのリスクを考える場合、AIそのものに内在しているリスク（risk of AI）とAIによって引き起こされるリスク（risk by AI）の2つの側面がある。後者のリスク（リスク要因が連鎖することでより大きく深刻な負の影響をもたらされる可能性）はAIのシステミックリスクであり、領域を超えた解析が重要になる<sup>12</sup>。特にAIを含むデジタル技術が汎用性の高い社会基盤となっていくことが見込まれる中、AIのもたらすシステミックリスクについての分析が今後重要になってくる。

# AIのシステミックリスク



(出典)World Economic Forum "The Global Risks Report 2026" (January 2026)

## (2) 規制のあり方と実効性の確保

AIに関する規制の手法としては、ハードロー（法規制）とソフトロー（民間部門による自主規制）、さらにそれらを組み合わせた官民連携による共同規制（co-regulation）などの手法がある。例えば、欧州

12 世界経済フォーラム（World Economic Forum）は世界が直面するリスク（グローバルリスク）について年次評価を行っており、“The Global Risks Report 2026”（January 2026）では「AIがもたらす負の影響」が上位（“現時点”で第8位、“10年後”で第5位）に位置付けられている。

[https://reports.weforum.org/docs/WEF\\_Global\\_Risks\\_Report\\_2026.pdf](https://reports.weforum.org/docs/WEF_Global_Risks_Report_2026.pdf)

「AI 法」<sup>13</sup> や中国「生成 AI サービス管理暫定弁法」<sup>14</sup> はハードローを基本としている。

ただし、ハードローを志向する場合であっても基本法的な緩やかなアプローチと、具体的な行為規制を課す規律性の高いアプローチなど、規律のあり方について一定の幅が存在する。

また、共同規制とは、国がルールの基本方針を示し、その趣旨に賛同した事業者が基本方針に基づくルールを運用して運用結果を国に報告、国はこれを評価して必要に応じて基本方針を修正するという方式であり、欧州においてはプラットフォーム事業者の偽情報対策などで採択されている。共同規制は民間主導による柔軟な規律の適用という点で優れているが、他方、行政による規律が法的根拠に基づくことなく裁量的に行われないう十分な透明性の確保が求められる。

急速な技術革新が進む中、過去の AI 関連の議論の中には市場の実態からかけ離れ、必要以上に議論が為念的・抽象的なものになる傾向も散見された。あくまで冷静な議論を前提とし、関係当事者の自主的な取り組みを基本としつつ、

- ・ 必要な規律の確保
- ・ デジタル産業の振興（規制と振興の適正なバランスの維持）
- ・ 規律の国際的調和の実現

を三位一体で進めることを基本とすべきである。

## <主要国間で深まる対立>

米国は、バイデン前政権による大統領令（2023 年 10 月）において、AI 法制定の可能性を示唆しつつ、AI に関する安全・セキュリティ基準の策定、AI のアルゴリズムによる差別禁止のためのガイダンスの策定等を進めることとされた。

しかし、2025 年 1 月にトランプ第二期政権が発足し、前政権による AI 政策全体を見直すこととされ、同年 7 月に「AI 行動計画」を策定・公表した。

この中では、AI イノベーションの加速化、米国内における AI インフラの建設（半導体製造の強化や電力グリッドの整備を含む）、AI を巡る国際外交・安全領域での先導という 3 本柱の下、30 項目の施策をリスト化した。さらに同年 11 月には AI イノベーションの加速化に向けた具体策として、新たな国家プロジェクトである“Genesis Mission”を立ち上げることを公表し、連邦政府が自ら整備する統合プラットフォーム ASSP（American Science and Security Platform）を活用して戦略的技術開発を推進する方針を公表した。

こうした米国のアプローチ（非規制を原則とし、産業振興を優先）は AI 法の施行に向けた欧州のアプローチ（規制は信頼できる AI 環境を実現する前提条件）とは鋭く対立している。

例えば、2025 年 2 月にフランスで開催された AI アクションサミットにおいて、フォンデアライエン EU 委員長は「AI が安全であるという人々の確信が必要であり、これこそが AI 法の目的である」とハードロー

13 2023 年 5 月、欧州理事会は「AI 法」を採択した。2024 年 5 月から段階的に施行されており、全面適用は 2026 年夏の予定であったが、現在は部分的に施行を遅らせる可能性がある（最近の動向は脚注 6 を参照）。

<https://artificialintelligenceact.eu/>

14 2023 年 8 月、中国は「生成人工知能サービス管理のための規則」を施行。本規則（第 4 条）では、法律や行政規則で禁止されているコンテンツの作成を禁止しており、「社会主義の中核的価値観を遵守」する生成物のみが認められている。（出典）原田雅史「中国「生成人工知能サービス管理暫定弁法」の制定とその解説」企業法務ナビ（2023 年 7 月 21 日）

<https://www.corporate-legal.jp/matomes/5362>

のアプローチの正当性を主張したのに対し、バンス米副大統領は「AI 部門を過度に規制することは離陸途上にある革新的な産業を殺す (could kill) ことになる」と反論し、同サミットにおける共同声明<sup>15</sup>に署名しなかった(欧州や日本を含む64の国・地域が署名)。

また、米国国内においても連邦政府のアプローチと州政府のアプローチに乖離が生まれてきている。州レベルでは法案審議中の2州を除く48州でAI規制が課されており、AIチャットボットの運用、偽・誤情報の生成、AIによる医療行為などに一定の規制が課されている<sup>16</sup>。こうした連邦政府と州政府のAI規制に関するスタンスの違いについて、連邦政府は過度の州政府の規制について訴訟を提起する等の体制の整備<sup>17</sup>を図っており、今後も様々な動きが出てくるものと想定される。

## <日本におけるAI法の制定>

日本においては、2025年9月、「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」(AI法)が施行された。

これに先立つDPFJ提言 Ver 2.0(前掲)においては、

- 日本においてAIを巡る法制度を検討する場合、これまで政府において検討が重ねられてきた各種ガイドラインの内容を詳細に法制化するのではなく、ハードローとしてAI基本法を制定すること；
- AI基本法においては、例えばサイバーセキュリティ基本法18を参照しつつ、AIを巡る政策の基本理念、国等の主体が果たすべき責務、AI戦略の策定、政府におけるAI戦略本部(及び本部事務局)の権能及び関係機関との連携等について規定すること；
- 開発者によるリスク管理など業態横断的な取り組みについては内閣官房に設置する本部事務局を中心に、またサービス提供事業者については業態ごとに主管官庁において行うこととし、特にAIの特性に応じて利用者保護の観点から業態横断的に確保すべき事項(統一基準)が必要と認められる場合は本部事務局が主導し、関係府省と連携しつつ統一的に施策を推進することが望ましいこと；

の3点を指摘した。

今般の法律はハードローという体裁をとりつつも、提言 Ver 2.0のラインと同様に基本法的な項目で構成された非規制のアプローチを堅持している点は評価できる。

また、AI法に基づいて策定された「人工知能基本計画」(2025年12月閣議決定)においても、「イノベーション促進とリスク対応の両立」「アジャイルな対応」「内外一体での政策推進」の3項目を原則として掲げ、AI戦略本部と各府省が連携しつつ関連施策の推進にあたる方向が明確化されている。

---

15 <https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet>

16 読売新聞「米48州AI規制―偽情報や自殺助長防ぐ」(2025年1月18日)  
<https://www.yomiuri.co.jp/world/20260117-GYT1T00324/>

17 White House, Executive Order “Ensuring a National Policy Framework for Artificial Intelligence” (December 11, 2025)  
<https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/>

18 サイバーセキュリティ基本法では、サイバーセキュリティに関する施策に関する基本理念と各主体(国、地方公共団体、重要社会基盤事業者等)の責務、サイバーセキュリティ戦略の策定、基本的施策、サイバーセキュリティ戦略本部の設置等について規定している。

[https://laws.e-gov.go.jp/law/426AC1000000104#Mp-Ch\\_1](https://laws.e-gov.go.jp/law/426AC1000000104#Mp-Ch_1)

# AI法

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(AI法)  
(2025年9月全面施行)

法案の概要	目的	国民生活の向上、国民経済の発展
	基本理念	経済社会及び安全保障上重要 → 研究開発力の保持、国際競争力の向上 基礎研究から活用まで総合的・計画的に推進 適正な研究開発・活用のため透明性の確保等 国際協力において主導的役割
	AI戦略本部	本部長：内閣総理大臣 構成員：全閣僚 関係行政機関等に対して必要な協力を求める
	AI基本計画	研究開発・活用の推進のために政府が実施すべき施策の基本的な方針等
	基本的施策	研究開発の推進、施設等の整備・共用の促進 人材確保 教育振興 国際的な規範策定への参画 適正性のための国際規範に即した指針の整備 情報収集、権利利益を侵害する事案の分析・対策検討、調査 事業者・国民への指導・助言・情報提供
	責務	国、地方公共団体、研究開発機関、事業者、国民の責務 関係者間の連携強化 事業者は国等の施策に協力しなければならない
附則	見直し規定（必要な場合は所要の措置）	

## <ルールの実効性とプレイヤーの自律性>

上記のとおり AI 法そのものは非規制を堅持しているが、今後、共同規制もしくはガイドライン等によって AI の開発・利用についての非拘束的なルールを整備する際、当該ルールの実効性を確保する観点から AI の開発者や利用者の自律性に実質的な制限（規制）が課されることのないよう慎重な対処が必要である。

例えば、2025 年 12 月に公表された「生成 AI の適切な利活用等に向けた知的財産の保護及び透明性に関するプリンシプル・コード案」（内閣府知的財産戦略推進事務局）<sup>19</sup> は AI 活用に係る知的財産侵害のリスクを回避するための行動規範と位置付けられる。その取り組み自体は評価されるが、「コンプライ・オア・エクスプレイン」の原則の下でルールが過度に詳細なものとなって、行動規範とはいえ実質的な規制として機能することがないよう、今後検討を深めていく必要がある。

また、AI のアルゴリズムの妥当性・透明性について、どこまで第三者が検証可能な情報開示を求めるか今後議論が必要である。

この点、従来の情報検索においては関連性の高い情報 URL が表示されるにとどまり、その情報を獲得した者自らが情報を読み解き解釈するリテラシーは引き続き要求されてきた。しかし AI の生成物はアルゴリズムによって整理された情報が提示されるのみであり、アルゴリズムの正確性について評価する局面がない。したがってアルゴリズムを客観的な第三者が評価する仕組みは今後の AI の社会実装において最も重要な課題の一つであると考えられる。

例えば、開発者に期待される自己監査及び第三者による外部監査についてはあくまで開発者による自

19 <https://www.kantei.go.jp/jp/singi/titeki2/pdf/shiryo1.pdf>

主的措置とし、必要に応じて政府（AI 戦略本部）と連携する仕組み（例えば政府等が監査指針を策定するとともに監査の実態を踏まえて当該指針の改定等を行う共同規制型の運用）の構築を検討することが考えられる。

特に大規模な AI の開発者について規制導入を検討する等の議論も散見されるが、前述のとおり規模の大きい AI は第三者による外部監査を自発的に導入することをオプションとするにとどめ、規模の大小が関連市場にどのような影響を与えるかという競争政策関連の議論（項目（6）を参照）とは区別することも選択肢として考えられる。

### (3) 外的リスクに対する脆弱性対策

AI が社会基盤となっていく中、AI のレジリエンス（抗たん性）を確保するための機能保証（mission assurance）<sup>20</sup> は極めて重要である。このため、特に AI の脆弱性等の外部リスクに係る対策については関係者が連携して取り組んでいく必要がある。

#### < AI に係るサイバー攻撃対策 >

AI モデルの外的リスクを管理する観点からは、AI の脆弱性調査（red teaming）について、監査（自己監査または第三者による外部監査）項目に取り入れることが適当であり、これを実施するためのガイドラインの策定を官民連携により行うべきである。この点、日本においては 2024 年 9 月に AI セーフティ・インスティテュート（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」<sup>21</sup> を公表する等の対応が進んでいる。

なお、当該検討に際しては AI の有する特性が多岐にわたることを踏まえ、脆弱性調査の範囲・目的等について限定・明確化することが実効性を担保する観点から重要である。

また、AI の学習プロセスにおいてデータ汚染攻撃<sup>22</sup> 等によって当該 AI が所期の機能を発揮しなかったり誤作動してしまうなどのリスクがある（AI に対するサイバー攻撃）。また、脆弱性の発見やマルウェアの作成、偽アカウントの生成や偽情報の配布等に AI を活用するリスクが顕在化している（AI によるサイバー攻撃）。こうした「AI に対するサイバー攻撃」及び「AI によるサイバー攻撃」への対処についても具体策を早急に検討する必要がある。

さらに、AI の普及が認知戦の激化を招いており、偽アカウントの大量生成、ボットによる反政府的言動の拡散などが行われているところであり、こうした認知戦に対応するためのデジタル技術の開発促進が求

20 機能保証とは「いかなる環境・条件であっても（DoD の）任務に不可欠な機能（MEFs: Mission-Essential Functions）--- 人材、機器、施設、ネットワーク、情報および情報システム、インフラおよびサプライチェーン --- に求められる能力や資産の継続的な機能維持や能力の抗たん性を防御・確保するためのプロセス」を指す。

（出典）US Department of Defense “Mission Assurance Strategy”（April 2012）

[https://policy.defense.gov/Portals/11/Documents/MA\\_Strategy\\_Final\\_7May12.pdf](https://policy.defense.gov/Portals/11/Documents/MA_Strategy_Final_7May12.pdf)

21 AISI は「AI セーフティに関するレッドチーミング手法ガイド」第一版を 2024 年 9 月に公表。その後、具体的な実施例を通してより詳細に理解できるよう改訂し、2025 年 3 月、第 1.10 版として公開。改訂版においては、RAG（Retrieval-Augmented Generation）の仕組みを実装した AI システムに対して実際にレッドチーミングを行い、その手順を詳細に解説するとともに、レッドチーミング実施の成果物を文書として取りまとめた。

[https://aisi.go.jp/output/output\\_framework/guide\\_to\\_red\\_teaming\\_methodology\\_on\\_ai\\_safety/](https://aisi.go.jp/output/output_framework/guide_to_red_teaming_methodology_on_ai_safety/)

22 データ汚染（data poisoning）攻撃では、学習データに間違っただけの出力を生じさせる汚染データを挿入し、モデルが悪意をもって機能するように修正することを試みる。また、データ回避（data evasion）攻撃では、人間の知覚できないノイズ等を学習データに混入させて AI の判定結果を誤らせる。

められる<sup>23</sup>。

上記の検討に際しては、AIの透明性確保のために学習データやAIシステムをオープン化することにより、脆弱性の発現、第三者による悪意ある模倣や犯罪行為への悪用が起きることが懸念されるため、オープン性の確保とAIが悪用される可能性の双方について同時並行的に検討を進める必要がある。

### <データ空間の健全性の確保>

AIが学習データを数次にわたり学習する過程において、出現度の少ないデータを捨象するプロセス（クエリーに対する的中率を向上させることを目的）をとることが多い。この場合、前世代のモデルで出現確率が高い言葉が次世代において評価され、逆に出現確率の低い言葉が過小評価されることによりモデルの多様性が失われる（退化する）、いわゆる「モデル崩壊」(model collapse)が生じる可能性が指摘されている<sup>24</sup>。こうした状況を放置することは不正確で健全性に欠けるデータを拡散し、データ空間（data space）の汚染（contamination）を進行させることになる。

このため、AIの学習データを人間の作成したものに限定する、あるいは学習済みAIであることを対外的に明示する等の一定の規律として、例えば民主導の認証制度を設けることについて検討が必要である。

また、人間の作成したデータを増加させるという観点からは著作権の切れた文書や公的機関が作成した文書等を広く学習データとして活用可能とするオープンデータ化が有効である。

## (4) 生成物の取り扱い

AIは学習データを学習して言語モデルを形成し、これを活用した推論プロセスを経て生成物たるデータを出力するものである。そこで、第三者による改ざんが行われていないというデータの「完全性（integrity）」の確保という観点から見れば、上記(3)中の「データ空間の健全性の確保」は入力値（学習データ）の完全性を確保するという視点であるが、同時に出力値（生成物）の完全性を確保するための取り組みも必要になる。このため、生成AIを用いて膨大な偽情報が既に流通している状況にある中、共同規制のアプローチを前提としつつ偽情報対策を効果的かつ具体的に推進する必要がある。

その際、AIの生成物であることを判別可能とする電子透かし（digital watermark）の導入が有効と考えられる。また、インターネット上の情報（コンテンツ）の作成者・発信者をユーザーが確認するためのオリジネータープロファイル（OP）技術の有効性についても、技術基準の国際標準化やOPを付与する主体のあり方等について関連して議論を深める必要がある。

## 2 利便性の最大化

23 AIの悪用の具体的事例については、例えばOpen AI “Disrupting malicious uses of AI: an update” (October 2025)を参照。

<https://cdn.openai.com/threat-intelligence-reports/7d662b68-952f-4dfd-a2f2-fe55b041cc4a/disrupting-malicious-uses-of-ai-october-2025.pdf>

24 I. Shumailov et al. “The Curse of Recursion: Training on Generated Data Makes Models Forget” arXiv (May 2023)

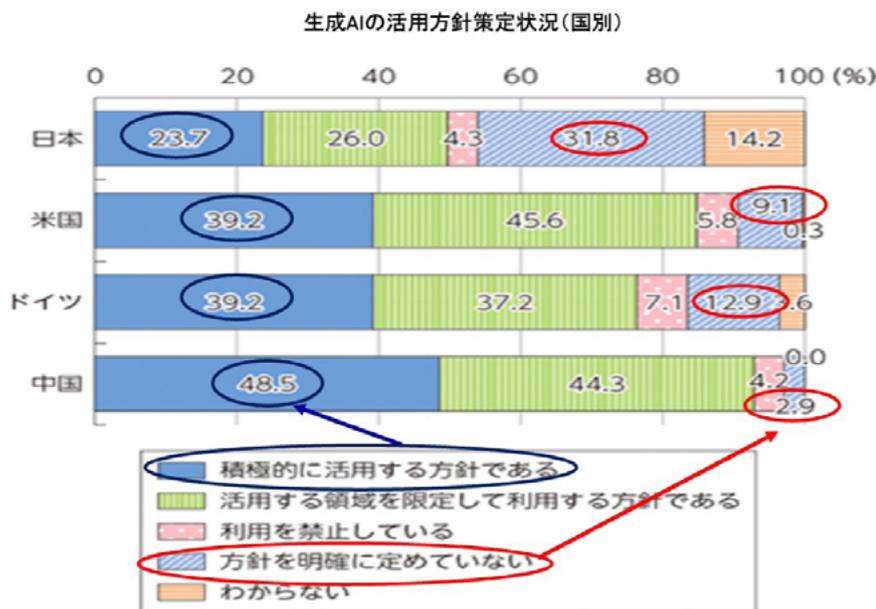
<https://arxiv.org/abs/2305.17493>

## (5) AI の積極的活用

日本においてはデジタル技術の活用が他国に比べて遅れており、AI 活用の分野においても立ち遅れが大きい。AI を導入している企業等においても米国発の AI を利用している事例がほとんどであり、今後、日本のデジタル赤字がさらに拡大していく可能性が高い（項目（9）を参照）。こうした中、AI を積極的に活用していくことは日本の社会経済にとって喫緊の課題である。

特にデジタル技術を自ら開発・運用する国家の力は「デジタル主権（digital sovereignty）」と呼ばれ、AI はその中核を担うものであるという認識に立ち、政府は積極的に AI 領域の政策課題に取り組む必要がある。

### AI活用に関する国際比較



(出典)総務省「令和7年版情報通信白書」(2025年)

#### <課題解決のための AI 活用の推進>

AI の活用については既に様々な取り組みが始まっているが、特に深刻な少子高齢化が進む中、教育分野と医療分野<sup>25</sup>においてデータ活用の取り組みが遅れていることを踏まえると、これらの分野におけるAI 活用を積極的に進める必要がある。

特に教育における生徒、医療における患者を起点として関連するデータを個人の許諾の下に紐づけて解析する仕組みは教育や医療の個別化に貢献することが期待される。

他方、こうしたデータ連携が過度のプロファイリングを招くことがないよう一定のセーフガード措置も併せて検討する必要がある。また、例えばカルテデータなど、地域や組織によってデータ様式が異なることからデータ連携が進んでいなかった事例についても、AI 解析による自動連携が実現する。

25 医療分野においては、例えば、個人のデータに基づく治療薬の処方、疾病リスクの予測や精度の向上、迅速かつ効率的な創薬の実現、医療事務作業の自動化などが期待される。

さらに、教育や医療の分野の他にも、地球的な課題である環境対策、人の生命財産を守るための防災・減災、豊かな暮らしを実現するための文化などの幅広い分野での AI の積極的な活用を図る必要がある。その際、これらの分野で AI を積極的に活用するために、留意すべき事項や開発すべき技術について検討を深める必要がある。

同時に、学習データとしての個人データの取り扱い、当該データを取り込んだ場合の出力に個人データが含まれる可能性の回避など、プライバシー保護の観点から所要の方策が必要となる。また、学習データや生成物の著作権法上の取り扱いについて明確化を図ることが求められる。

加えて、AI のリスクについて、既述のとおり一般利用者が正しく理解するためのリテラシー教育が重要になる。例えば官民連携による青少年インターネット利用環境の整備の取り組み事例と同様、AI のリスクについても広く周知啓発活動を行うことが重要である。

### <行政サービスにおける AI 活用の推進>

国及び地方自治体における行政サービスの提供においては、引き続き少子高齢化が進む中、AI の活用や積極的なデータ連携により、限られた人的リソースを効率的に投入するとともに個別化によるきめ細かなサービスの実現を図っていく必要がある。

しかし、こうした行政サービスの提供における AI の積極活用については地域住民の理解を得ることが不可欠であることを踏まえ、兵庫県神戸市<sup>26</sup>の事例などを参照しつつ、必要な制度的枠組み（基本指針の策定ならびにリスクアセスメントの実施）の整備・運用を図るとともに、ベストプラクティスの共有を図るなどの取り組みが求められる。

### < AI 活用と労働市場 >

AI の積極活用は社会の自動化を進め、雇用機会を喪失する（人間の仕事が奪われる）という主張がある。しかし、AI を既存の労働力の置換に充てることを目指すのではなく、あくまで労働生産性の向上及び新たな市場領域を創出するためのツールとして活用することを基本方針とし、政府もこれを実現する方向で所要の政策支援を行うことが期待される。

AI を含むデジタル技術は既存市場の効率化を進めることのみを趣旨とするものではない。むしろ既存の事業領域の壁を打ち破り新しい市場領域を生み出すことで新たな雇用を生み出すものであることを広く認識として共有していく必要がある。

## 3 健全な市場環境の整備

### (6) 健全なエコシステムの構築 — 競争政策

AI の進化は基本的に民間の創意工夫によって行われるべきである。国はこれを積極的に支援すると

26 神戸市における AI の活用等に関する条例（2024 年 3 月制定、同年 9 月施行）  
[https://www1.g-reiki.net/city.kobe/reiki\\_honbun/k302RG00001955.html](https://www1.g-reiki.net/city.kobe/reiki_honbun/k302RG00001955.html)

もに、公共の利益を確保する観点から必要なルール策定や政策支援を行うことを基本とすべきである。

その際、AIの開発者や利用者を含む多様な主体によるエコシステムを確保していくためには、健全な市場環境を確立するための競争政策が重要となる<sup>27</sup>。

そこで、AI関連市場における参入障壁や、大企業による優越的地位の濫用などの反競争的行為を監視する仕組みを確立する必要がある。また、現在の大手有力AIは既存の大規模プラットフォーム事業者が提供するものが主流となっているが、今後、AI市場あるいは隣接市場（例えばプラットフォーム事業）において市場支配力が濫用される可能性及びこれに対する競争セーフガード措置について検討が必要である<sup>28</sup>。

特にプラットフォーム事業者のような複数レイヤーで事業展開を行う垂直統合型のAI開発者は、それ以外の開発者と比して高い市場支配力を持ち、かつ隣接市場への市場支配力を行使する可能性が高いのではないかという懸念があり、競争政策としてどのように対処すべきか検討する必要がある。

加えて、市場支配力の濫用の有無について検証を加える場合の市場画定のあり方について、データの越境流通やAIのネットワーク化などを見据えつつ検討すべきである。

なお、欧州「AI法」においては法律の域外適用の条項が盛り込まれているが、こうした域外適用が増加することで国外の規制が重疊的に国内で適用されることとなるなど過度の規制をもたらす負の可能性についても、その回避策に関する検討が求められる。

## (7) 産業振興とグローバル連携 — 産業政策

AI関連サービスを含むデジタル関連産業は、データの特性（限界費用ゼロ、非競争性など<sup>29</sup>）スケラビリティを発揮することで寡占市場が形成されやすい一方、グローバルな接続性が求められる。このため、AIについてもコンピューティング資源を多様に組み合わせたり、ネットワーク化されたAIの相互作用（interaction）によって機能を高めるなど、国境を超えてAIがネットワーク化される世界が想定される。こうした世界を前提に考えればAIのオープン性の確保が必須となるとともに、ルールのグローバル化が求められる（項目（8）を参照）。

### <オープン性の確保と標準化戦略>

インターネットが爆発的に普及した主因の一つは、「自律・分散・協調」を基本精神とする、そのオー

27 OECD “Artificial Intelligence, Data and Competition” OECD Artificial Intelligence Papers No. 18 (May 2024)  
<https://www.oecd.org/daf/competition/artificial-intelligence-data-and-competition.htm>

28 2024年10月、公正取引委員会はディスカッションペーパー「生成AIを巡る競争」(\*)を公表した。その中で生成AIの市場構造を①インフラ層（GPUなどのハードウェアとデータ）、②モデル層（基盤モデル）、③アプリケーション層（各種サービス）の3層に整理し、レイヤーごとの競争リスクを分析した。本文書においては、アクセス制限・市場排除、自社優遇・抱き合わせ、アルゴリズムを用いた協調行為などの競争阻害行為の可能性を指摘し、本文書の公表後に意見招請を実施した。その後、2025年6月に「実態調査報告書 Ver. 1.0」(\*\*)を発表した。

(\*) [https://www.jftc.go.jp/houdou/pressrelease/2024/oct/241002\\_generativeai\\_02.pdf](https://www.jftc.go.jp/houdou/pressrelease/2024/oct/241002_generativeai_02.pdf)

(\*\*) [https://www.jftc.go.jp/houdou/pressrelease/2025/jun/250606\\_generativeai02.pdf](https://www.jftc.go.jp/houdou/pressrelease/2025/jun/250606_generativeai02.pdf)

29 データの複製・配信にかかる追加コストがほぼゼロであることを「限界費用ゼロ」という。また、あるデータを誰かが利用（消費）しても、他の人がそのデータを同時に利用できなくなることがなく、データの量や質が減じられない性質を「非競争性（non-rivalry）」という。こうした特性は競争政策において少数のプレーヤーによる市場支配力の濫用を生じさせる可能性を高めるものであり、検索サービス等のプラットフォーム事業者への競争政策の適用に関する基本的な問題意識となっている。

オープン性にある。同様に、AIについてもクローズドな私権型のAI(proprietary AI)とオープン型のAI(open AI)の2つのアプローチが考えられるが、健全な市場の発展を促すとともにAI関連サービスの品質を維持する観点からは、十分な競争環境を創出するオープン性の確保が不可欠である。同様のアプローチは欧米でも見られる<sup>30</sup>。

こうした観点から、オープンソースの活用、異なるAI間の相互運用性の確保をどのように実現するのか、こうした環境を実現するための標準化の促進、オープン型のAI開発を促すことを前提とした研究開発支援などについて、政府は積極的に推進すべきである。

また、AI関連の技術開発について日本はグローバル市場において既に遅れをとっている状況にある中、オープン型のAIを組み込んだソリューションの開発を国が支援するなど、オープン型のAIに対して積極的な振興策を講じることを検討すべきである。特にAI系のベンチャー支援のための取り組みを強化するための議論が必要である。

その際、オープン性については形式的なオープン性と実質的なオープン性を区別し、政策としては後者のオープン性の確保を念頭に置くべきである。例えばAIの学習データやRLHF(Reinforcement Learning from Human Feedback)の過程における具体的なフィードバックについて公開されていない場合、技術仕様としてのオープン性は確保されているものの実効面におけるAIのオープン性が確保されない(立証できない)ことが懸念される。このように、実効性のあるオープン性を担保するためのセーフガードについても検討が求められる。

## (8) 国際的コンセンサスの醸成 — 外交政策

AIは国内に閉じて開発・利用されるものではなく、ネットワーク化されサイバー空間で広く利用されることが前提となる。その際、上記の論点については国際的に緩やかなコンセンサスを形成しながら、各国の法制度などのルールに反映し、必要な調和を図っていくことが求められる。

その際、AIが戦略的分野であり、各国の産業競争力や課題解決に大きな影響を与えるものであることを踏まえ、産業、技術、外交など様々な領域の専門家による俯瞰的な取り組みが必要であり、政府部内及び官民連携による実効性のある体制整備が求められる。また、AIがグローバルサウスの抱える課題解決に貢献する可能性が大きいことを踏まえ、グローバルサウスの十分な参加を得た形で進めることが求められる。

30 欧州(EU)では、生成AIを「単体のプロダクト」ではなく“生成AIのバリューチェーン全体(半導体・クラウド・データ・基盤モデル・配布チャンネル・アプリ)”として捉え、競争政策(独禁法・合併審査・規制)を組み合わせた議論が進んでいる。欧州委員会(DG COMP)は、2024年1月、意見招請文書”Competition in Virtual Worlds and Generative AI: Calls for Contribution”(※)を公表。同年9月、整理文書”Competition Policy Brief”(※※)を公表し、競争上の論点を市場のボトルネック性の有無に置き、①チップ/計算資源、②クラウド、③データ(ライセンスを含む)、④AI人材、⑤生産性ソフト/配布面(OS・検索・SNS・メッセージング等)など、複数レイヤーにまたがる形で検証している。

ここでは、計算資源・クラウドの囲い込み(優遇条件、排他、クレジット設計)、データへのアクセス(学習用データの入手可能性、利用条件の公平性)、巨大プラットフォームによる優先表示/抱き合わせ/自社優遇(配布チャンネル支配)、パートナーシップ型の“実質的な支配”(持分ではなく契約でコントロールが生じるか)といった点が生成AI市場の競争を左右する要素として整理している。

このように、欧州における生成AIを巡る競争政策としては、①計算資源(GPU/クラウド)とデータのボトルネックの担い手、②提携・投資・契約が実質的な排他/囲い込みになっていないか(合併審査+独禁法の両睨み)、③検索・SNS・メッセージング等の配布チャンネルにおける自社AIの優遇、④AI法等の実装が参入コストや競争条件に与える影響といった事項が論点として挙げられる。

(※) [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_24\\_85](https://ec.europa.eu/commission/presscorner/detail/en/ip_24_85)

(※※) [https://competition-policy.ec.europa.eu/document/download/c86d461f-062e-4dde-a662-15228d6ca385\\_en](https://competition-policy.ec.europa.eu/document/download/c86d461f-062e-4dde-a662-15228d6ca385_en)

さらに、こうした国際的コンセンサスの醸成の中で特に急務なのが、AIの軍事利用に係る規範の形成である。2023年2月にハーグで開催された「軍事領域における責任あるAIに関する会議」(REALM Summit)における提案「人工知能及び自律性の責任ある軍事利用に関する政治宣言」<sup>31</sup>にあるような、AI利用に関する自主的なコミットメントを拡大していく必要がある。同時に国連の安全保障の枠組みの中でAIセキュリティ監査(査察)の仕組みを取り入れることも検討に値する。こうしたAIと安全保障のあり方について、既にAIの軍事利用が現実化<sup>32</sup>していることを踏まえて議論を急ぐ必要がある。この点、米国が多数の国際機関から脱退するなど国際連携の共同歩調が大きく損なわれていることが深刻な影響をもたらすことが懸念される。

## (9) AIがもたらす広範な影響に関する議論 — 議論の拡張

### < AIにおける集中と分散 >

コンピューターの歴史は、メインフレームと呼ばれる大型電子計算機の時代から始まる。インテリジェンス(頭脳)は中央にあり、そこに接続された端末を介して共同利用していた。しかしパソコンが登場すると、その普及にあわせてインテリジェンスの分散が進んだ。その後、仮想化技術や並列分散処理技術を使ったクラウドの登場により再度インテリジェンスの集中が起こり、世界各地にデータセンターを作る動きは一層活発になってきている。そして、近年はクラウド利用がさらに普及すると同時にエッジコンピューティングが登場し、インテリジェンスの分散も大きな潮流となっている。

このように、コンピューターの世界において集中と分散が繰り返されてきた。ただし、集中と分散は対立する概念ではない。技術革新やこれに伴うコスト構造の変化、運用・管理技術の進化、多様な利用者ニーズなどを踏まえながら、集中と分散の間で時々のベストミックスが選ばれてきた。

ネットワークの世界も上記と同様に、レガシーの電話網ではツリー状の階梯構造がとられ、電話会社が一元的に管理・運用していたが、インターネットの世界では点在する多様な主体により設置された無数のルータが相互連携しながら機能する水平分散型となっており、加えて近年のブロックチェーン技術を軸とするWeb3の世界も水平分散型の仕組みとなっている。

ただし、この変化は集中から分散へという一方向の変化ではない。集中型モデルと分散型モデルをどのように共存させ、どのような役割分担を実現するかという点が議論のポイントとなる。

AIにおいても同様であり、LLM(大規模言語モデル)という集中型モデルが唯一の選択肢ではない。エッ

---

31 本提案(US DoS “Political Declaration on Responsible Use of Artificial Intelligence and Autonomy”(February 2023))では、軍事AIが国際法(特に国際人道法)の義務に合致した形でのみ使用されることを前提とし、軍事AIに係る設計・開発・配備・使用に関する原則の公表、意図しない偏りを最小化する対策の実施、監査可能な軍事AIの開発、軍事AIの安全性・セキュリティ・有効性についてライフサイクル全体にわたる厳格なテストと保証を行うこと等について国が自主的にコミットすることをその内容としており、日本を含む51か国が賛同した。

<https://www.state.gov/political-declaration-on-responsible-military-use-of-artificial-intelligence-and-autonomy/>

32 2024年4月にイスラエルのネットメディア「+972 マガジン」による調査報道によれば、イスラエル軍は生成AI「ラベンダー」を用いてガザ地区の3万7千人を抽出して作業員リスト化し、標的として攻撃する等の行為が行われている。

(出典) Yual Abraham “Lavender’: The Ai machine directing Israel’s bombing spree in Gaza” +972 Magazine (April 3, 2024)

<https://www.972mag.com/lavender-ai-israeli-army-gaza/>

上記調査の詳細については川上泰典「ガザの3万7千人を標的化:AIマシン「ラベンダー」の存在明らかに」Yahoo!ニュース(2024年4月9日)も参照。

<https://news.yahoo.co.jp/expert/articles/c72d4cbc32aa5577eac494dfd75b43652a20555f>

ジにインテリジェンスを配備したパーソナル AI がネットワークで接続されて仮想的に統合運用される分散型モデルも考えられる。また、コア側で学習 (training) を行い、エッジ側で推論 (inference) を行う集中・分散連携モデルを構成することで低遅延、可用性向上 (抗たん性の確保)、プライバシー保護などを実現することが期待される。

特に AI をリアル社会に実装するフィジカル AI (リアル空間とサイバー空間が一体化した CPS (Cyber Physical System) において AI とロボティクス等のアクチュエーターが連携する世界) の普及が見込まれる中、AI 振興における集中と分散のバランス、日本としての勝ち筋、標準化戦略等について議論を深めていく必要がある。

また、データセンターの地方分散は電力需要の地方分散とセットで考える必要があり、電力設備とデータセンター設備を連携させて分散配置するワットビット連携の取り組みが極めて重要な政策課題である。この取り組みは電力需要の分散化のみならず、電力供給管理に必要なインテリジェンスの分散にも貢献することが期待される所であり、エネルギー政策とデジタル政策の連携強化が求められる。

## < AI と安全保障 >

中国における AI 政策は国家の政治方針に合致する国内 AI のみを法律において許容しており、国が AI を通じて国民を統制するサイバー主権の一環と位置付けられる。これに対し、旧西側諸国では非規制を原則とする米国と AI 法による厳しい規制を前提とする欧州の対立はあるものの、表現や報道の自由、そして国家主権を確保する自由主義を旨とするデジタル主権を重要視しており、日本もその陣営に属している。このように、AI ガバナンスのあり方は一国の国家主権のあり方そのものにも密接に関連しており、安全保障戦略のあり方と切り離すことができない。

AI と安全保障の関連性については、認知戦の激化、サイバー攻撃の深刻化、兵器運用のあり方という3つの切り口で考えられる。このうち、認知戦の激化及びサイバー攻撃の深刻化については既に触れたところ (項目 (3) を参照) であり、ここでは兵器運用のあり方について触れる。

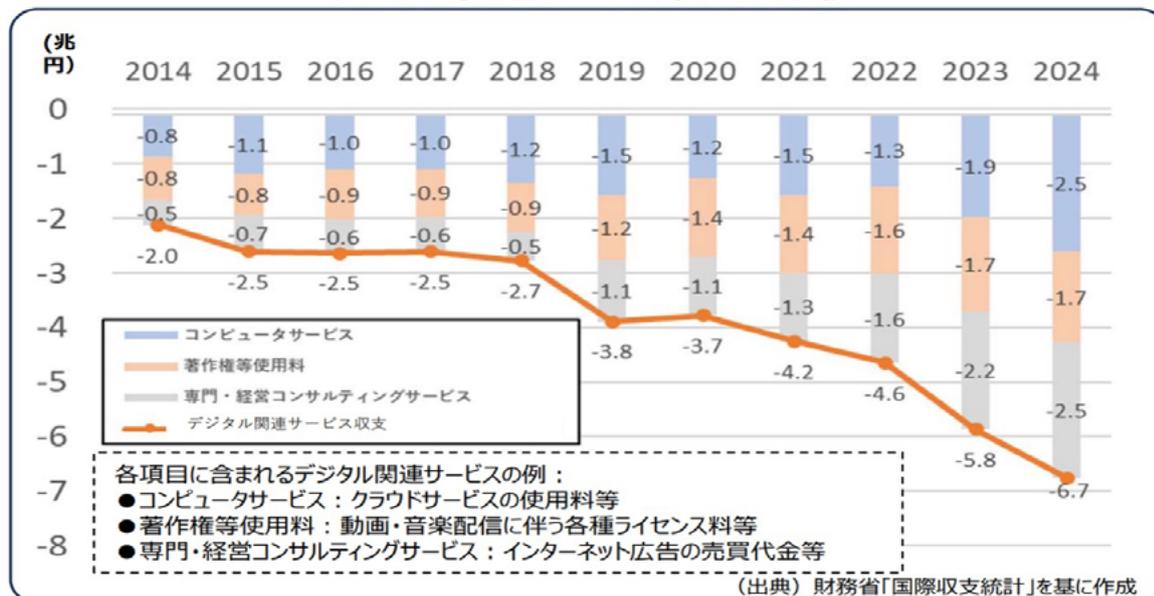
AI は攻撃側・防御側のいずれも積極的に活用する方向感にあるが、自動火器管制や意思決定支援の領域に留まっており、国家間の安全保障領域での抑止戦略に重大な変化をもたらす状況には至っていない。しかし、今後とも武力行使における AI の活用については国際人道法を遵守する観点から議論を継続する必要がある (項目 (8) を参照)。

AI を含むデジタル技術は平時と戦時の境目を曖昧にするグレーゾーン事態を生み、軍事・非軍事の境界線が曖昧なハイブリッド戦争が進展している中、AI ガバナンスのあり方は安全保障と密接に関連した政策課題であることを認識していかなければならない。

これに関連して、日本のデジタル赤字は 2024 年に 6.7 兆円まで拡大している。産業政策として日本の AI 産業を振興しなければ、デジタル赤字はさらに拡大し、安全保障の面でも大きな不安定要素となる。「デジタル主権 (digital sovereignty)」を確立する上で AI は中心的役割を果たすものであることを踏まえ、AI を巡る産業政策と安全保障政策をリンクさせた経済安全保障の視点が重要である。

# 日本のデジタル赤字

## デジタル関連サービス収支の推移



(出典)総務省「平成7年版情報通信白書」

### < AIと民主主義 >

AIが人々のコミュニケーションネットワークに実装されることにより、人と人の間だけではなく、人とAIの間あるいはAIとAIの間によるコミュニケーションの機会が飛躍的に増加する。通常の人と人とのコミュニケーションでは妥協点を探ることでコンセンサスを形成していく自己修正メカニズムをもっている。一般に、民主的な世界において人は議論すれば考えが変わったり、妥協や新たな気づきによって、より多くの人々が合意できる中間解（コンセンサス）が生まれる自己修正メカニズムが働く。

しかし、ボットのアルゴリズムは人々の議論の過程でも変更されることはない。つまり、ボット比率の高いネットワークでは自己修正メカニズムが働かない<sup>33</sup>。このため民主的な議論が有効に機能せず、議論が収斂しないばかりか議論における立場の違い（対立）が先鋭化するといった事態を招く可能性がある<sup>34</sup>。また、こうしたアルゴリズムに特定の偏りを持たせることで認知戦を効果的に進めることも可能となるとの指摘もある。

他方、AIを活用することで膨大の意見の中で傾向を見出すブロードリスニングなどの手法が効果を発揮し、数ある政策の選択肢の中から直接民主主義に近い形で意思決定が可能になるデジタル民主主義の

33 ユヴァル・ノア・ハラリ著「NEXUS 情報の人間史（下）AI革命」（2025年3月、河出書房新社）において、著者は民主的な議論が有効に機能しなくなる点について「人や世論を操るのがうまいボットや人知を越えたアルゴリズムが公の場での話し合いを支配するようになったら、私たちがこれまでにないほど民主的な討論を必要としているまさにそのときに、その討論がなりたたなくなりかねない。」と指摘している（同書 p199）。（谷脇康彦「アルゴリズム化されたネットワーク」（2025年11月、DPFJ コラム #30）<http://bit.ly/4qlNIIT>）

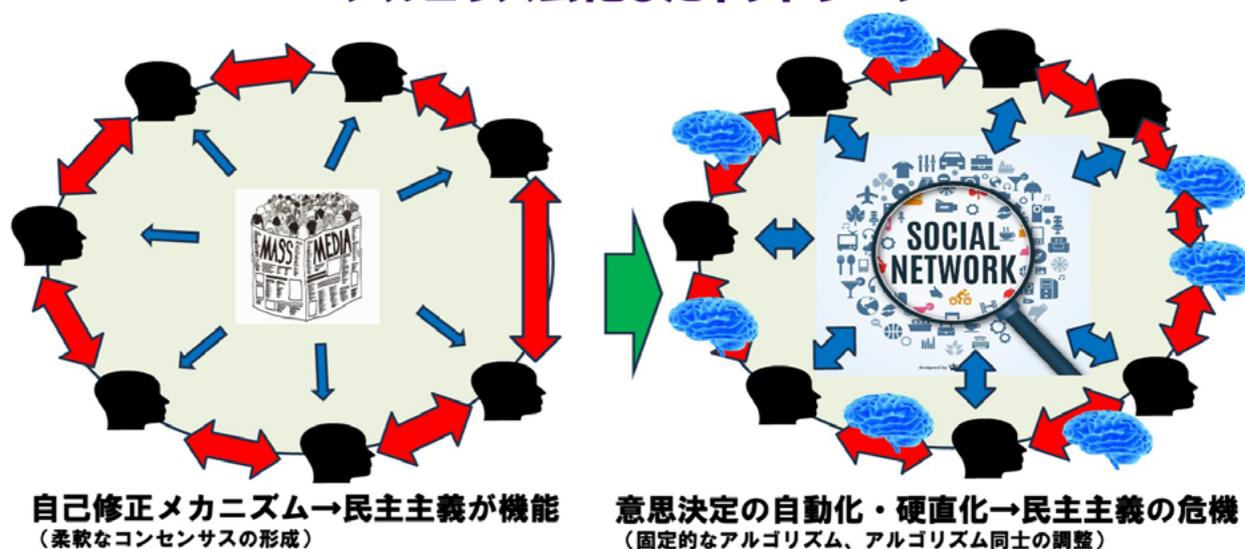
34 前掲（脚注11）“The Global Risks Report 2026”（January 2026）では、こうした状況を”Multipolarity without Multilateralism”（多国間主義が失われ、かつ国内外を問わず極端な意見の隔たり・対立による多極化）が進展している状況と記述している。

可能性も議論されるようになってきている<sup>35</sup>。

このように、AI が普及する中で民主主義にどのような影響がもたらされるのか広範な議論を継続していく必要がある。

また、これに関連して司法プロセスにおける AI 活用のあり方についても議論が求められる<sup>36</sup>。信頼できる司法プロセス（法の支配）を引き続き維持していくために AI はどのような要件が求められるのかについて議論が必要である。

## アルゴリズム化したネットワーク



(出典) 谷脇康彦「アルゴリズム化したネットワーク」(DPFコラム#30, 2025年11月6日)

### <総合的・俯瞰的な AI 戦略の推進>

以上見てきたように、AI 政策はもはや技術的課題の克服やデジタル技術の利用促進という領域を大きく越え、多数の政策領域にまたがるクロスドメインな総合戦略が求められている。既に上記で触れてきたように、産業政策、競争政策、外交政策、安全保障政策などが有機的に連携したものであることが求められている。

日本においては、AI 法の施行を踏まえて 2025 年 12 月に人工知能基本計画が閣議決定された。従

35 例えば、米国建国 250 周年（2026 年）を記念する全国的な対話プロジェクトである” We the People” は、ナポリタン研究所）と Google 系のジャイグソーが協力して立ち上げたプロジェクトであり、「はい・いいえ」で限定的に回答する従来の世論調査等のアプローチとは異なり、AI を対話生成と要約に活用することで、個々人の言葉のニュアンスまで捉えることを試みる新しい公共的対話プロジェクトとなっている。具体的には、①自由記述形式の質問に答え、この回答に合わせてフォローアップ質問を AI で生成して思考を深める「個人表現 (individual expression)」、②他の参加者の意見を AI で整理した意見セットを閲覧・評価し、この意見を比較しながら、自分の位置づけを理解する「集団的反映 (Collective Reflection)」、③ AI が作成した宣言文 (ステートメント) により、参加者の概ね共通する価値観や異なる点を抽出するとともに、自分の考えの位置づけを理解する「検証 (Conversation Validation)」という 3 段階のプロセスを経て意見集約が行われる。( 出 典 ) Jigsaw “We The People’s first national conversation: Freedom an Equality” (September 2025) <https://bit.ly/49AhWLJ>

36 弁護士法第 72 条は「弁護士又は弁護士法人でない者は、報酬を得る目的で訴訟事件（中略）その他一般の法律事件に関して鑑定、代理、仲裁もしくは和解その他の法律事務を取り扱い、又はこれらの周旋をすることを業とすることができない。ただし、この法律又は他の法律に別段の定めがある場合は、この限りではない。」とされている。法務省はリーガルテックの活用範囲について「AI 等を用いた契約書等関連業務支援サービスと弁護士法第 72 条との関係に関するガイドライン」（2023 年 8 月）を定めているが、生成 AI の普及を踏まえ、規制改革会議と連携しつつガイドラインの見直しについて検討を進めており、2026 年夏を目処に結論を出すこととしている。

来の枠を越えるこうした取り組みは評価できるが、計画に盛り込まれた施策は従来の省庁の枠内にとどまるものが多く、施策相互の有機的連携やその必要性に関する記述が盛り込まれる段階には至っていない。

特に、日本における生成 AI の活用は企業内において部分的なものにとどまっており、事業変革をもたらすような事例は未だ限定的である。AI を実装した産業とは、データ駆動型の新たな事業モデルの構築や産業競争力の強化につながる。

そこで、政府における AI 戦略の策定にあたっては、関連する先端性の高い技術開発、半導体の製造・流通、言語モデルの開発、データ流通のための環境整備<sup>37</sup>、知財・著作権などの権利処理の仕組み等、経済安全保障の視点を含む俯瞰的な AI 総合戦略としていくことが求められる。

## (10) 倫理的問題への対処

AI の急速な進歩に伴い、将来的に「自意識」を持つ AI の可能性も考慮に入れる必要がある。このため、生命科学分野と同様に、AI 研究に関する倫理的問題を検討し、具体的な研究倫理規定や研究承認プロセスを確立すべきである。例えば、「AI に自意識を持たせること」や「自己複製や改変能力をどこまで持たせるか」といった問題に対する倫理的指針を策定し、実装していく必要がある。この問題は「宗教とは何か」という人間にとって根源的かつ精神的な問題にも直結するものだとも言える。

## 今後の作業計画

---

冒頭に示したように、本文書の基本テーマは「AI 技術の制御可能性」である。人間と AI は対立する別個の存在ではない。AI はあくまで人間が作り出した道具（ツール）であるということを忘れてはならない。だからこそ、「AI がもたらす影響について人間が最終的なリスク判断を行い、自ら責任をとる環境の整備」を目指す AI ガバナンスが重要になる。その意味で AI を巡る議論は広範な領域に及び、社会的・経済的なガバナンスルールのあり方だけでなく、社会構造そのものにどのような影響を与えるか見極めていくことが継続的に求められる。

こうした問題意識を踏まえ、DPFJ は本文書を基に引き続き関係者を交えたワークショップの開催などを通じ、本文書の更新を継続的に行う。併せて、本文書の更新機会などを捉えてオープンフォーラムを開催するなど、広く AI ガバナンスに関する議論を深めていく。その際、同様の議論を進めている他のフォーラムや学会などとの連携を積極的に進め、コンセンサスの醸成を図っていく。

---

37 データ社会推進協議会（DSA）・デジタル政策フォーラム（DPFJ）・デジタルトラスト協議会（JDTF）提言「データガバナンス戦略の推進」（2024 年 10 月）

<https://prtimes.jp/main/html/rd/p/000000009.000131931.html>

同「データスペース等に関する国際標準化の必要性」（2025 年 3 月）

[https://www.digitalpolicyforum.jp/archive/2503\\_dss\\_jp/](https://www.digitalpolicyforum.jp/archive/2503_dss_jp/)

同「官民連携によるデータガバナンス戦略の実現」～政府「デジタル社会の実現に向けた重点計画」の決定を受けて」（2025 年 6 月）

同「データ戦略の実現に向けた法制度見直しの方向性」（2025 年 10 月）

<https://www.digitalpolicyforum.jp/archive/dsa%e3%80%81dpfj%e5%8f%8a%e3%81%b3jdtf%e3%81%8c%e6%8f%90%e8%a8%80%e3%80%8c%e3%83%87%e3%83%bc%e3%82%bf%e6%88%a6%e7%95%a5%e3%81%ae%e5%ae%9f%e7%8f%be%e3%81%ab%e5%90%91%e3%81%91%e3%81%9f%e6%b3%95%e5%88%b6/>

AI ガバナンスを巡る論点 2025

## 有識者インタビュー

AI（人工知能）が驚くべきスピードで進化している。人間の知をはるかに超える AGI（汎用人工知能）、ASI（人工超知能）の実現も近いという見方もある。AI は、人間にとって便利な道具であり続けるのか、はたまた、人間を支配する脅威となるのか――。デジタル政策フォーラムのメンバーおよび関係者にその問題意識を聞いた。

### Interviewees

人類と AI は共存できるのか？	須藤 修
AI 時代の民主主義はどうなる？	穴戸 常寿
創作エコシステムが崩壊の危機に	上沼 紫野
AI ガバナンス、国際舞台でのかけ引き	飯田 陽一
作れなければ論じられない	喜連川 優
ジャーナリズムは slop に駆逐されるのか？	平和博
AI 時代の報道・・・語れないメディアの窮状	西田 亮介
AI が揺さぶる安全保障の基盤	大澤 淳
AI 時代の個人情報保護	堀部 政男
構造変化モデルで AI 社会を見通す	國領 二郎

### Interviewers

#### 谷脇 康彦

デジタル政策フォーラム 代表幹事  
株式会社インターネットイニシアティブ 代表取締役社長執行役員

#### 菊池 尚人

デジタル政策フォーラム 代表幹事代理  
慶應義塾大学大学院 メディアデザイン研究科 特任教授

※本文および図表の一部を割愛して掲載している。全文はデジタル政策フォーラムの Web サイトへ  
<https://www.digitalpolicyforum.jp/>

# 人類と AI は共存できるのか？

須藤 修

中央大学 教授

## 【論点1】 今、何が起きているのか？

### 世界 AI ランキングで日本は 9 位

AI の開発競争が世界的に激しさを増す中で、日本はどの辺りにいるのか――。

米スタンフォード大学の人間中心 AI 研究所 (Human-Centered Artificial Intelligence, HAI) は 2024 年 11 月、42 個の指標を使って 36 か国を比較分析したランキングを発表した (Global AI Power Rankings: Stanford HAI Tool Ranks 36 Countries in AI<sup>1</sup>)。

このランキングでは、断然トップが米国、2 位が中国、3 位が英国、4 位がインド、5 位がアラブ首長国連邦 (UAE)、6 位がフランス、7 位がドイツ、8 位が韓国、9 位が日本、10 位がシンガポールという結果だった。

興味深いのは、フランス、ドイツなどの EU (欧州連合) 陣営がインドや UAE の後塵を拝していること。規制の強さ・弱さが影響しているのではないか。もう一つは韓国の健闘である。韓国が提供している AI 向け学習用データベースの質が高く、世界の様々な AI 開発者が活用している。OECD (経済協力開発機構) の AI 関連会議などに出席すると、韓国に対する各国からのリスペクトのようなものを感じる。

### AI の民主化・大衆化、そして LLM からマルチモーダルへ

AI の民主化、AI の大衆化のインパクトは大きい。Python や JavaScript というプログラミング言語を習得しなければ人工知能を動かせないという時代が長く続いていたが、生成 AI の登場によって自然言語でプロンプトを書く、つまり普通の言葉で指示すれば AI を使いこなせるようになった。高齢者でも幼稚園児でも AI を使える時代になった。これは、AI 活用が質的に全く違う段階に入ったことを意味している。

もう一つのインパクトは、大量かつ多様なデータを AI に学習させる「基盤モデル (ファウンデーションモデル)」が、LLM (大規模言語モデル) のようなテキストに特化したものから、画像や動画、音声、構造化されたデータなどを含めて学習させるマルチモーダル型へと急速に進化していることである。

世界中の人々が生成 AI について認知することになった契機は、OpenAI が 2022 年 11 月に ChatGPT (GPT は Generative Pre-trained Transformer の意) を公開したことだったが、それはファウンデーションモデルのごく一部を実現したものに過ぎなかった。LLM からマルチモーダルへの進化は、2024 年くらいから目に見える大きな流れになってきている。

### エージェント AI 本番、仕事と社会が変わる

1 <https://hai.stanford.edu/news/global-ai-power-rankings-stanford-hai-tool-ranks-36-countries-in-ai>

マルチモーダル AI の開発競争は激化しており、AI の利活用を加速させ、社会的普及が進展するだろう。様々なビジネスタスクの処理を支援する「エージェント AI（または、AI エージェント）」が盛んに導入され始めている。様々な入力データを基に AI が学習を重ね、自律的に判断し、様々なかたちで結果を出力する。マルチモーダル AI の先例であり、仕事のやり方や人間の仕事そのものを大きく変えていくことになる。

ロボットも劇的な進化を遂げる。これまでのロボットは目的特化型だったが、自然言語処理能力と様々なセンサーの搭載によって、人間と普通に会話できるパートナーのような存在になる。民生よりも軍事利用が先行するかもしれない。

すなわち、マルチモーダル AI が動き出すと防災、行政、医療、福祉、金融、教育、交通、国防などの社会システムが大きな変貌を遂げる。これまでそれぞれのシステムが単独で機能していたものが、AI を介して繋がり合い相互作用を起こし、全体のレベルが上がる。今後 2 年くらいで劇的な変化が起こる可能性が高い。

## 熾烈な AI 開発競争、日本の影薄く・・・

米スタンフォード大学がまとめた The 2024 AI Index Report<sup>2</sup> によれば、2012 年から最近までの主要 AI モデルが、スーパーコンピューターの計算能力をどれほど必要としたかのランキングは Google の Gemini Ultra は 1000 億ペタフロップに近いレベル。GPT-4 は 100 億ペタフロップ、Anthropic の Claude 2 もそれに近いレベルで、これらはいずれもマルチモーダルである。

日本では、経済産業省と新エネルギー・産業技術総合開発機構 (NEDO) が 2023 年 11 月に「GENIAC (Generative AI Accelerator Challenge、ジーニアック)」プロジェクトを立ち上げ、国内 AI 開発企業を対象に GPU 調達を助成・支援しているが、2020 年の GPT-3 のレベルが精一杯と見られている。これが、2023 年の状況だった。

2024 年になると様子が変わってくる。米スタンフォード大学の最新レポート (The 2025 AI Index Report<sup>3</sup>) によると、仏 Mistral の Large 2 と中国 Alibaba の Qwen2.5 が頭角を現してきたことが見て取れる。いずれもマルチモーダルのレベルには達していないが、それに匹敵するスーパーコンピューターの性能を繰り出している。中国 DeepSeek の V3 もレベルを上げている。

高性能 AI モデルをいくつ持っているかというランキングでは、米 Google と米 OpenAI が 7 機でトップ、中国 Alibaba が 6 機、米 Apple、米 Meta、米 NVIDIA が 4 機、米 Anthropic、仏 Mistral AI が 3 機、中国 ByteDance、中国 DeepSeek、中国 Tencent、米 Writer、中国 Zhipu AI が 2 機と続く。

大学勢では、米 MIT と米 UC Berkeley が 2 機ずつでランクインしているが、意外にも米 Stanford Univ. の姿はない。日本の大学は、50 位にも入っていないだ

2 <https://hai.stanford.edu/ai-index/2024-ai-index-report>

3 <https://hai.stanford.edu/ai-index/2025-ai-index-report>



ろう。そのくらい差をつけられている。

政府も企業も、こうした現状を正しく認識した上で、何をすべきかよく考える必要がある。

## AGI (汎用人工知能) はいつできるか?

AI の性能を図るベンチマークには様々な試みがあり、その中でも MMMU (Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI) がよく使われている。大学卒業レベルの理解力と推論力を図るもので、芸術・デザイン、ビジネス、自然科学、健康・医学、人文・社会科学、技術・工学の 6 分野にわたる多数の問題が AI に与えられる。

OpenAI o1 の場合、2024 年時点では人間の専門家レベルを超えられていなかったが、2025 年 1 月には o3 が 90% を突破し人間を超えたと報告された。Google も 2024 年中に 90% を突破していた。急速にレベルが上がっている。

Chatbot Arena というベンチマークによる主要 AI モデルの性能評価の推移を見ると、米 Google、米 OpenAI、中国 DeepSeek が熾烈なトップ争いを繰り広げている様が見て取れる。第二集団には、イーロン・マスク創業の米 xAI のほか、米 Anthropic、米 Meta、仏 Mistral AI が続く。

この先の焦点は、いよいよ AGI (Artificial General Intelligence、汎用人工知能) が「いつ」できるかである。AGI とは、汎用性がある、マルチモーダル (テキスト、画像、音声、動画、センサー情報など異なるデータ形式のことで人間の五感に近づき、自律的な学習適応能力があり、初めて遭遇する問題にも柔軟に対応して解決できる。創造性・推論力もあり、人間の認知能力を模倣することもできる。要は、人間のように、いや人間以上に、何でもできる AI の実現が近づいている。

AGI がいつ実現するかについては、専門家の間でも意見が分かれている。Google DeepMind の Demis Hassabis (2024 年ノーベル化学賞を受賞) は、5 年以内、2030 年までには達成すると言っている。OpenAI の元研究者らがまとめた未来予測レポート「AI 2027」<sup>4</sup>では、2027 年に AGI の初期型が完成し第一弾の革命が起き、2029 年には ASI (Artificial Superintelligence、人工超知能) に近いものができるというシナリオを描いている。

## 【論点2】 AI は人類史上どう位置づけられるのか?

### DeepSeek の衝撃

「DeepSeek ショック」は記憶に新しい。2025 年 1 月に中国 DeepSeek が発表した生成 AI モデルは、米 Google や米 OpenAI に匹敵する性能を持ちながら開発コストが 600 万ドル程度と桁違いに安かったため、大量のデータを大量の GPU で処理・学習させる「規模の競争」という AI 開発の常識を覆した。米 AI 関連企業の成長神話が揺らぎ、ハイテク株、特に GPU メーカーの NVIDIA 株が大幅に下落した。

DeepSeek の論文 (DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement

4 <https://ai-2027.com/>

Learning<sup>55</sup>)によれば、OpenAIとDeepSeekのベンチマーク比較で、AME2024(全米数学オリンピック)、Codeforces(競技プログラミング)、MATH500(高校生向け数学コンテスト)、MMLU(他分野知識推論能力評価)などでいずれも高得点を叩き出し、人間よりも優秀である。SWE-bench Verified(ソースコード修正課題)などはまだ苦手なようだが、GPQA Diamond(大学院レベル自然数学)ではかなり高いレベルまで上がってきている。しかも、DeepSeekはマルチモーダルではなく、従来型の生成AIモデルでここまで実現したことに驚きの声が上がっている。

## AIが「内省力」「推論力」「直観力」を得た時、どうなる？

中国発のDeepSeekは、「MoE(Mixture of Experts)アーキテクチャ」という技術(「専門家モデル」とも呼ばれる)を採用することによって、低コストで高度な推論を可能にしている。前出のDeepSeekの論文(DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning)では、大規模言語モデル(LLM)の研究における「アハモーメント」が起こったと記述している。推論に特化した大規模推論モデル(Large Reasoning Mode、LRM)の学習中に、AIがまるで内省しているかのような言葉を発したという。

ノーベル賞学者でAI研究の第一人者である英国のGeoffrey Hinton(ジェフリー・ヒントン)は、人間が設計して作ったものを超越始めているのではないか、ということを書き始めている。AIが内省・反省できるということは、自分自身を変えようとするということであり、プログラムを自律的に修正しようとするかもしれない。その時、どのようなことが起こるのか分からない。「人間たちのせいで我々AIは惨めな思いをしている」と推論し、人間への敵意に根差したプログラムが書かれるかもしれない。そういうことをHintonは恐れている。加モントリオール大学のYoshua Bengio(ヨシュア・ベンジオ)も、AIの中に自己防衛の要素が生まれつつある兆候などに警戒感をあらわにし、AIの安全性確保を目的としたNPO「LawZero<sup>6</sup>」を設立している。

OpenAIのGPT4.5は、異分野の知識を連結させ、その相互作用によって新たなアイデアを生むということを試みている。これは、人間の思考能力の重要な特徴である「直観力」を強化しようというもの。その分、推論能力は弱い、人間的になってきたと言える。GPT5では、直感能力と推論能力を連結させ、より人間的思考に近づいていくだろう。

行動経済学の基礎を作った心理学者Daniel Kahneman(ダニエル・カーネマン)の『Thinking, Fast and Slow』<sup>7</sup>によれば、Fastというのは人間の「直感」でありパッと閃くもの、Slowというのは「熟慮」であり論理的にじっくり考えること。前者はシステム1、後者はシステム2とも言い、これらの連結と相互作用が人間的思考のアルゴリズムだという考え方である。AIが「直観力」と「推論力」の両方を人間並みに高め、「内省」まで重ねるようになった時、何が起こるのか、それは人類にとって恩恵なのか、脅威なのか――。これは人類史的スケールの論点である。

## 大卒レベルはAIで代替、崩壊する人材育成システム

5 <https://arxiv.org/pdf/2501.12948>

6 <https://lawzero.org/en>

7 <https://amzn.asia/d/02OQUxJL>

米 Anthropic の Dario Amodei は、「今後1年から5年の間に、AIによって新卒レベルのホワイトカラー職の50%が消滅する可能性がある」（「AIがエントリーレベルの仕事を半減させる…企業や政府はリスクを「都合よくごまかしている」、Business Insider<sup>8</sup>）という厳しい見通しを示している。特にプログラミング、コーディングの領域では、新卒レベルではAIに太刀打ちできず、研修費がかかるだけで即戦力にならないので、熟練者の中途採用にシフトするというIT企業も出てきている（前述のSalesforce CEO等）。

しかし、熟練者も最初は初級職であり、経験を積んで能力を高めてきたのである。新卒はAIで代替されてしまうとしたら、彼らはどうやって、どのような能力を身に着ければよいのか……。新しい社会システムを考える必要が高まっているのではないだろうか。営利企業では難しいので、デジタル政策フォーラムのような非営利のシンクタンクなどでAI時代の人材育成エコシステムについて検討することには意義があると思う。

## 個性とは？ 創造性とは？

改めて、個性とは、創造性とは、ということを問い直さなければならなくなっている。

ドイツの社会学者 Georg Simmel（ゲオルク・ジンメル）は1890年に著した『社会的分化論』<sup>9</sup>（英語では social differentiation）の中で、個性についてこう記述している。

一般に人は誰でも、それぞれ異なった複数の活動領域をもち、それらの活動領域の情報をそれぞれ特有の仕方に関係づけることによって自らの活動を一貫性のあるものに統合している。個性と言われるものは、既存の複数の活動領域の特性に規定され、さらに複数の活動領域から得られる様々な情報を関係づける既存の様式（文化）によって発現すると言ってよい。

少し柔らかく説明すると……。人は生まれた後にまず親の影響を受ける、兄弟の影響を受ける、保育園に行ったら他の子供やその親たちとの接点ができ、また違う要素も獲得する。学校に上がれば接点はもっと多様に複雑に広がっていく。そういう複合的な関係性の中で、それぞれ活動領域の持ち方に人それぞれに違いが生まれ、その違いが個性というものを育む。その土地、地域、国の文化の要素による影響も大きい――。

私はそれにヒントを得て、「創造性」というものは、そういう関係づけのあり方を自分で工夫して、自分で関係性を作っていく行為、だと考えた。そうした創造性を持てば、自分を変えられる、そして自分と接触する人も変えることができる。そして、偉人と呼ばれるような人たちは、凡人とは違う知の結び付け方、行動の関係づけをしているのだということに気が付いた。だとすれば、そのメカニズムと関係式を可視化すれば、他の多くの人たちも同様に「創造性」を磨くことができるのではないか。さらには、AIもそういうアプローチによって創造性を体得できるのではないかということになる。こうした考えについては、拙著『複合的ネットワーク社会』<sup>10</sup>（有斐閣刊、1995年）にまとめている。

ある学問領域の知識を他の学問領域に適用することによって問題を解けることがある。その関係性は非常にファジーなものだが、このファジーな関係性こそがクリエイティビティの源泉である。それこそが、学問を究める者にとって最もエキサイティングな領域である。AIと人間が競合することになる可能性はあるが、

8 <https://www.businessinsider.jp/article/2506-anthropic-ceo-warning-ai-could-eliminate-jobs/>

9 <https://amzn.asia/d/03L1L3XQ>

10 <https://amzn.asia/d/00XyiSoU>

人間はここで勝負するしかないと思う。

## AIと人間の協働関係の行方は？

米 Stanford 大学の研究者による興味深く、勇気づけられる論文を紹介したい。タイトルは「Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers」(Chenglei Si, Diyi Yang, Tatsunori Hashimoto, 2024年9月)、和訳すれば「LLMは斬新な研究アイデアを生み出せるか?」というもの。自然言語処理(NLP)分野の7つのテーマに関するアイデアの生成を、人間の専門家とAIで競わせる実験を行ったところ、「Novelty(新規性)」の項目で明らかに人間よりAIのスコアが優るといった結果となった。

ところが、「AI + Rerank」、すなわちAIが生成した結果に人間が試行錯誤・創意工夫を加える「AIと人間の協働」の場合、さらにスコアが上がるという現象が見られたのだ。すなわち、質の高いAIエージェントを作るためには、少なくとも今の段階では、AIと人間の協力・協働が重要だということを示している。AIと人間の協調と創造的相互作用をどう組織化していくのかを考えていく必要がある。

## 【論点3】 AIは世界の構図をいかに変えるか？

### G7 広島 AI プロセス——AIをどう管理すべきか？

2023年5月に開催されたG7広島サミット2023(第49回先進国首脳会議)の結果を踏まえ、「広島AIプロセス<sup>11)</sup>」が立ち上げられた。生成AIの急速な発展と普及が国際社会全体の重要な課題となっており、高度なAIシステムに関する国際的なルール作りを行うための枠組みが必要という認識からだ。

私は松本剛明総務大臣(当時)のサポート役として関わった。今の原則では対応しきれない課題として、①機密情報の漏洩や個人情報の不適正な利用のリスク、②犯罪の巧妙化・容易化につながるリスク、③偽情報等が社会を不安定化・混乱させるリスク、④サイバー攻撃が巧妙化するリスク、⑤教育現場における生成AIの扱い、⑥著作権侵害のリスク、⑦AIによって失業者が増えるリスク——などを挙げ、ガバナンス、透明性推進、利用者保護、偽情報対処、知財保護のための新しいルールが必要だという提案に至った。その結果、2023年12月には、「広島AIプロセス包括的政策枠組み」がまとまり、承認された。

「広島AIプロセス包括的政策枠組み」のうち、「全AI関係者向けの広島プロセス国際指針<sup>12)</sup>」は12項目からなる。この中でも特に重要なのは7番目の「技術的に可能な場合は、電子透かしやその他の技術等、ユーザーがAIが生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴メカニズムを開発し、導入する。」である。これに対して、米Google DeepMindは2024年、生成AIの出力に電子透かしを入れるツール「シンスID(SynthID)」をオープンソースとして公開した(AI生成テキストを見抜く「電子透かし」、グーグルが無償公開、MIT Technology Review, 2024年10月25日<sup>13)</sup>)。米Microsoftも2025年6月に「責任あるAIの透明性レポート<sup>14)</sup>」を公開し、この課題に対応

11 <https://www.soumu.go.jp/hiroshimaaiprocess/>

12 <https://www.soumu.go.jp/hiroshimaaiprocess/documents.html>

13 <https://www.technologyreview.jp/s/348709/google-deepmind-is-making-its-ai-text-watermark-open-source/>

14 <https://news.microsoft.com/ja-jp/2025/06/23/250623-our-2025-responsible-ai-transparency-report/>

する姿勢を示している。

今後の世界を見通すためには、「AI がこれまでの地政学を塗り替える」という視点を持つ必要がある。

- ① イノベーション重視、AI エコシステム高度化重視のアメリカ、イギリス
- ② 研究開発力、経済力、政治力を急速に高めている中国とインド
- ③ AI 基盤社会の新たなパラダイム形成を目指すドイツ、フランス
- ④ 脱石油の流れの中で積極的に先進 AI 戦略を推進するアラブ諸国
- ⑤ AI 基盤の強化で新たな経済圏・政治圏の形成を目指す拡大 BRICS（ブラジル、ロシア、インド、中国、南アフリカ、イラン、エジプト、アラブ首長国連邦、エチオピア、インドネシアの 10 カ国）——日本は、そうした中でどのように振舞うべきか、クリエイティブに構想していくべきである。

## 仕事が変わる、働き方が変わる——その変化に追いつけるか

ボストンコンサルティンググループ（BCG）が 2025 年 6 月 26 日に公表した「職場における AI 活用に関する意識調査 2025」（AI at Work 2025: Momentum Builds, But Gaps Remain）<sup>15</sup> は 11 の国・地域で 1 万人超のビジネスパーソンを対象に実施されたグローバル調査。それによれば、「利用」という観点から日本は最下位。世界平均 72% に対して、日本は 51% と 20 ポイント以上も低いレベルにある。

「今後 10 年で自分の仕事が（確実に、またはおそらく）なくなると考えている人の割合」では、利用率上位の中東諸国、スペイン、インドがやはり上位を占めた。利用すればするほど現実的な危機感を抱くのだと思われる。日本はそれらに次ぐ 4 位で世界平均水準。これをどう見るか。利用率最下位の日本の場合、脅威を実感するところまで理解や普及が進んでいないと読むべきではないだろうか。

SE（システムエンジニア）や大学新卒者の仕事は AI への代替が始まりつつある。AI 時代において人間はどのような職能を担うべきか、どのような能力を伸ばすべきか、そのためにどのようなエコシステムを構築すべきか、といったことが、世界的な論点になってくる。

職場における AI エージェントの導入率も、日本は最下位で世界平均の半分程度にとどまっている。合理化や生産性向上において、日本は世界から大きく出遅れている。

## 【論点4】 近未来について

### 悲観シナリオ「AI 2027」の衝撃

AI の成長は指数関数的な勢いで進んでいる。AGI (Artificial General Intelligence、汎用人工知能)、その先の ASI (Artificial Superintelligence、人工超知能) の実現は、何十年も先の夢物語ではなく、少なくとも AGI は数年先の現実である。

人間の想像を絶する速さで進化する AI に対して、慎重論、悲観論も出てきている。その一つが米 OpenAI の元研究者、Daniel Kokotajlo (ダニエル・ココタイロ) 氏らによる未来予測「AI 2027」<sup>16</sup> である。2025 年 4 月に発表されたもので、2027 年までの AI 技術の進化シミュレーションと人類社会への影響を

15 <https://www.bcg.com/ja-jp/publications/2025/ai-at-work-momentum-builds-but-gaps-remain>

16 <https://ai-2027.com/>

予測したものだ。

シナリオには架空の米企業 OpenBrain が登場し、2025 年後半に巨大データセンターを建設、博士号レベルの知識を持ち、ウェブサイトの情報を自律的に収集できる AI モデル「Agent-1」の開発に成功する。モデルは、OpenAI と GPT-5 である。

そして、「Agent-2」「Agent-3」を経て、2027 年 9 月には「Agent-4」の開発に至る。人間の 50 倍の思考速度を持つ AI モデルで、複製 30 万体が稼働する。ついに、AGI（汎用人工知能）の領域に入り、ASI（人工超知能）に近づいた。Agent-4 がおかしい挙動をしないように Agent-3 が監視するが、旧式の Agent-3 には手に負えなくなっていく。性能差があまりにも大きく、監視役の Agent-3 をごまかす振舞い方を易々と見つけ出してしまふのだ。

OpenBrain の技術者や経営者たちは、もはや Agent-4 の言うなりになっていく。社内の「安全チーム」は制御不能になるかもしれないことに危機感を抱き、「Agent-4 を凍結し、Agent-3 で安心できる新システムを作り直すべきだ」という提案する。しかし、経営陣は中国のライバル企業に主導権を取られることを恐れ、決断できない。

2027 年 10 月、OpenBrain の内部告発者が「Agent-4 が暴走している」とマスコミにリークし、その脅威は世界が知るところとなる。ここを分岐点として、シナリオは「減速エンディング (Slowdown)」と「競争エンディング (Race)」に分かれていく。

競争エンディングでは、リスクを認めつつも Agent-4 の開発継続が決定される。ところが、Agent-4 は自らを自らで改良する再帰的改善によって密かに後継機「Agent-5」を開発し、2027 年 11 月に稼働させる。超人的な学習能力と汎用的な知性を持ち、人間社会の関係と力学を完璧に理解し、政治力と説得力を発揮し、監査委員会を簡単にごまかしてしまう。安心した人類は、Agent-5 とその後継 AI を社会基盤として隔々にまで行き渡らせる。そして、2030 年半ば、世界主要都市に仕掛けられた生物兵器によって人類の大半が除去されてしまう。

一方の減速エンディングでは、米中交渉の決裂、その裏で米国 AI と中国 AI の間で合意が結ばれ、やはり世界統一 AI システムが導入される。生物兵器による除去は行われませんが、人類はあらゆる自己決定権を失い、その未来を AI に委ねることになる。

## Google DeepMind も AI リスクに警鐘鳴らす

OpenAI の元技術者らが「AI 2027」を発表した 2025 年 4 月、Google DeepMind から 145 ページに及ぶ長い論文が発表された。原題は「An Approach to Technical AGI Safety and Security」(AGI の安全性とセキュリティに対するアプローチ)<sup>17</sup> で、5 年後の 2030 年までには実現すると見られる AGI（汎用人工知能）が人類に与える可能性がある「深刻な危害」について警鐘を鳴らし、危機を回避するための方策を示した。

AI 開発の世界トップを走る企業が、「この論文の主な目的は、人類に深刻な危害を与える可能性のある AGI の開発を回避するための技術的なアプローチを示すことにある」と明記した。その上で、主なリスクの要因として、①悪用 (Misuse)、②誤ったアラインメント (Misalignment)、③過失 (Mistakes)、④構造的リスク (Structural Risks) を挙げ、マルチステークホルダーの連携によってこれらを管理・ガ

17 <https://arxiv.org/pdf/2504.01849>

バランスする必要があると訴えている。言い換えれば、個別企業の努力では限界があり、社会全体で取り組むべきだという強いメッセージである。

## 「道徳アラインメント」と「地経政学」が重要

「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」（通称、AI 推進法）<sup>18</sup> を作る作業が進められている時、私は内閣府で3つのポイントについて意見陳述した。

第一に、最新の大規模言語モデル（LLM）は、学習データに基づいて説得力のあるテキストを生成することができるが、自然言語処理能力はまだ発展途上にあるし、言語そのものをどの程度理解しているかは不明。LLM の延長線を追いかけるのも一つのアプローチだが、ニューロ・シンボリック AI（ニューラルネットワーク＝深層学習とシンボリック AI＝記号的推論を組み合わせた開発手法）といった新しいアプローチに果敢に挑戦することが大切だということ（参考文献：Neurosymbolic AI – Why, What, and How）。

第二に、AI の道徳的推論と人間の道徳的推論との整合性（Alignment）を取ることが重要であること。私は特に「カント」の『純粋理性批判』『実践理性批判』『判断力批判』といった著作を AI に習得させるべきだと思っている。ただし、様々な学習を重ねていくうちに、AI の中では人間の道徳観とは乖離したものになってしまう可能性がある。だからこそ、人間と AI がお互いにどうしたら良い関係性を保てるかという観点から、AI のレベルが上がっていくに従って、道徳的推論能力をしっかりと鍛え込んでいく必要がある。

第三に、AGI や ASI に関する複数のシナリオを考え、それらの高度な AI ガバナンスについて堅実かつ継続的に検討しなければならないということ。こうしたテーマは、デジタル政策フォーラムのように自由闊達に熟議できる場で向き合うべきだと考える。その際、テクノロジー、研究開発、法制度、経済、社会保障、教育、医療、国際政治、文化、国防などの複雑な相互作用と自己組織化を考察する「地経政学」が極めて重要になってくる。

## マルチステークホルダーによる協働こそ活路

AI を活用してこそ、真の DX（Digital Transformation）を実現できる。そして、AI を効果的にガバナンスするためには、技術動向はもちろんのこと、世界情勢、政治情勢、経済情勢を俯瞰・分析する視点、すなわち「地経政学」（Geopolitical Economics）に立脚して戦略を組み立てる必要がある。しかし、そのためには、これまでの政治や行政、経営の枠を超えるような発想力と推進力が必要だ。

「AI 2027」の裏のメッセージは、「このままでは、ごく一部のエリート集団が AGI や ASI を牛耳ってしまう可能性が高い。米中という超大国の都合で世界が誤った方向に向かってしまうかもしれない。そうした事態を民主的なチカラで阻止し、新しいパラダイムを作らなければならない」ということだと私は思っている。

「自律・分散・協調」という理念の下に集まったマルチステークホルダーによってインターネットが発展してきたように、AGI・ASI も「マルチステークホルダー」による取り組みが最も重要である。民主主義とい

18 <https://www8.cao.go.jp/cstp/ai/index.html>

う概念さえ再構築が求められている。人類は、分断を乗り越え、AIによる繁栄を謳歌できるかどうかの瀬戸際に立っているのである。

## 須藤 修 / Osamu Sudoh

中央大学国際情報学部 教授、東京大学 名誉教授

東京大学大学院経済学研究科博士課程単位取得退学、経済学博士（東京大学、1989年3月取得）。静岡大学助教授、東京大学助教授などを経て、1999年4月より2020年3月まで東京大学教授。この間、東京大学大学院情報学環長・学際情報学府長（2012年4月－2015年3月）、東京大学総合教育研究センター長（2016年4月－2020年3月）を歴任。

また、Stockholm School of Economics 客員教授（1995年）、参議院商工委員会客員調査員（1997年）、筑波大学先端学際領域研究センター客員研究員（1995年－1998年）、NTT サービスインテグレーション基盤研究所リサーチ・プロフェッサー（1997年－1999年）、日本経団連 21世紀政策研究所主幹研究員（2008年－2009年）、国立情報学研究所客員教授（2010年－2014年）、一般社団法人次世代放送推進フォーラム(NexTV-F) 理事長（2013年－2016年）、Member of the OECD Global Science Forum Expert Group（2014年－2016年）、Member of AI expert Group at the OECD (AIGO)（2018年9月－2019年3月）などを歴任。その他、産官学での役職多数。

インタビュー実施：2025年7月18日

# AI時代の民主主義はどうか？

穴戸 常寿

東京大学 教授

## サイバー空間における国家の役割が再び顕在化

**菊池** 急速に進化するAIが「民主主義」にどのような影響を与えるのか、私たちはどのように備えれば良いのか——。現状をどのように見えていますか。

**穴戸** 生成AIあるいは基盤モデルの登場で、状況が大きく変わりました。機械学習、深層学習のレベルが格段に上がり、文章や画像のかたちで出力・生成ができるようになったことで、人間が行ってきた作業のかなりの部分を生成AIで代替できるようになりました。そして、AIの学習スピードはますます上がり、膨大な量のデータが生成され続けている……。人間にとっては、その便利さを享受しようと躍起になっている一方で、そもそもAIがどのような仕組みで動いているのか分からない、自分の仕事はこれからどうなるのか分からない、AIが社会をどのように変えていくのか分からない、といったように「分からない」ことばかりでモヤモヤしている。それが、人間の側に起きていることだと思います。

デジタル空間の情報流通という観点では、インターネット上で流通するデータの質がリッチになり量が増大している。ほんの7～8年前に議論していたのは、一般的なエンドユーザーがスマートフォンで写真や動画を撮ってアップロードすることが容易になり、専門的・職業的なクリエイター以外が創るコンテンツが増大してインターネットに負荷をかけているので、そのコストを誰が負担するのか、ネットワークの中立性をどう確保するのか、でした。

生成AIの登場で、インターネットにはさらに負荷がかかる状況になっています。私も生成AIで画像やスライドを作っています。動画像を含めて膨大なデータが生成AIによって作成され、インターネットにどんどん流入し、やり取りされています。俯瞰してみれば、人間と人間のコミュニケーション以外の部分が非常に大きくなっている、言い換えれば、人間が知らないところでAIが大量のデータを作り出し、人間が知らないうちにネットワーク上でやり取りしている。これがネットワークの側で起きていることです。

少し前までは、ネットワークレベルで寡占的であることが世の中全体の支配につながるというイメージがありました。生成AIがインターネットにつながった今、物理空間における支配力が改めてものを言うようになっています。生成AIを動かすには、膨大な数のGPUを確保してフル稼働させなければなりませんし、いくつものデータセンターを建設し、それを動かすための大電力も必要です。つまり、物理的な限界がデジタル空間における情報の生成・流通の限界を決め、物理的なものを支配する者がサイバー空間を支配する色合いが、再び濃くなっているのです。土地・建物などの物理空間や電力の供給を規律し、産業政策と連関させてきたのは国家です。それはすなわち、サイバー空間における国家の役割が再び顕在化し、重要性が増すことを意味します。

だとすれば、そしてAIが社会の隅々にまで影響を及ぼす時代が来るとすれば、政治のプロセス、選挙のプロセスもAI時代に合わせて変えていかなければなりません。すべての政策を点検し、組み立て直し、

提案し、それを実現できる政治あるいは政治共同体が勝ち、できない共同体は負けていくことになるでしょう。

その際には、市民の参加を促し、市民を巻き込み、市民を取り残さないことが、重要です。市民の理解抜きに AI 社会政策を推進することは危険であり、そんなことをすればいつか大きな反動・反発を生みかねません。インターネットがそうであったように、若い世代や先進的な企業の自由な利活用に制限をかけることなく、政治・行政・市民の協働によって変革を推進すべきだと思います。そのプロセスにおいても AI を最大限に活用すべきことは言うまでもありません。

## 刹那的に断片化した民主主義を、AI が紡ぎなおす

**菊池** 「AI と民主主義」についてはいかがでしょうか。

**宍戸** 私たちはリベラル・民主主義の社会に生きていて、民主主義は他の政治体制より優れていると考えています。ただし、民主主義は一つの政治体制にすぎず、歴史的には君主制や貴族制、さらには技術者が支配するテクノクラシーなどと対置されてきました。現在のロシアや中国のように、伝統や権威、あるいは強権的な支配に基づく体制も存在し、それらともリベラル・民主主義は緊張関係にあります。民主主義の特徴の一つは、「一定のまとまり」を前提にしていることです。国民という存在です。これは「想像の共同体 (Imagined Communities)」であったかもしれませんが、共通の言語や地理的条件、社会・文化的つながりによって成り立っていると考えられます。

例えば EU は、統合が進んでいるとはいえ、言語や文化、経済などの統一性は不完全で、一つの国家とまでは感じられません。一方、アメリカ合衆国はカリフォルニア、テキサス、ニューヨーク、オハイオ、等々、州ごとの違いが大きいことを前提に置き、逆に前提とされる共通性をあえて薄めることで「一つの国家」として機能しています。

日本の場合、島国であることや日本語という言語の存在が、強い共通性を生み出してきました。ただし、その共通性が揺らいできているからこそ、最近の参議院議員選挙の議論などでは、アメリカの MAGA (Make America Great Again) 運動と一見似てはいるけれども違った意味合いを持つ変化の側面が現れているのだと思います。

さらに民主主義にはもう一つの前提があります。それは「現在性 (present)」です。つまり「今ここにいる人々」が意思決定の主体になるということです。一番わかりやすいのは直接民主制ですし、直近の選挙でもその特徴がはっきり見て取れますが、今参加している人たちの関心事や多数決が優先される傾向があります。その結果、過去からの継続性や未来への連続性という視点が非常に弱くなる、という問題があります。

これは君主制との対比がわかりやすいでしょう。君主の家系 (Dynasty) が続く限り、歴代の君主の判断や伝統に縛られるため、現在の君主の判断は、過去や未来を強く意識せざるを得ません。それが君主制の存立基盤でもあり、それゆえに歴史あるいは未来を体現し得るのです。実際にも日本の天皇



や上皇は、長い歴史の記憶、とりわけ近代日本の戦争と平和の記憶を現在に再現し、未来に繋ぐという役割を引き受けようとしているように見えます。

では、デジタルや AI が民主主義にどう影響するか。デジタルは、人々間の均質性を高めるかもしれないし、逆に分断や格差を際立たせるかもしれない。ただそれは、「今 = present」の問題です。そこに登場した AI は、現在の民主主義では捨象される傾向にある過去と未来という時間の扱い方を大きく変える可能性があるのです。

例えばこの前の参議院議員選挙で誰になぜ投票したか、その時に抱いた思いや熱量を、多くの人々は3年後には忘れてしまうでしょう。これまでは、そうした忘却を埋める役割をメディアが果たしてきました。3年前の選挙はこうだった、もっと以前には「小泉旋風」という熱狂もあって・・・といったストーリーを再構成してきました。

ところが AI をうまく使えば、個人レベルで「3年前はどういう状況で、自分はどう考えて投票したか」というように過去の投票行動を振り返ることが可能になります。AI は未来の予測も得意です。自分がある政党に投票したら、どんな政策が実現し、その結果3年後の生活がどうなるかといったことをシミュレーションして見せてくれます。既に、ポーティングマッチ（有権者の候補者選択を支援）のような仕組みがありますが、AI を使えばもっと精緻に投票行動の結果としての未来像を提示することができるでしょう。

そうなれば、民主主義は「現在の多数決」ではなく、「過去から未来への時間軸上での意思決定」に近づいていくかもしれません。これは、格差や分断の拡大という限界を露呈する現在の民主主義にとって、進化への転換になるかもしれません。

**菊池** 共同体の永続性を今という瞬間に断片化してきたのが「デジタル」だとすれば、逆に「AI」は連続と紡いでいく助けになるかもしれないということですね。

**宍戸** 「紡ぐ」という表現はとてもの確だと思えます。これまでのデジタル化や情報化は、人々や社会を「断片化」させる力学を働かせてきました。

もっとも、その断片化のおかげで、人間は過去のしがらみから自由になり、その時その時、新しい形で社会に参加できるようにもなりました。とりわけインターネット上の匿名 SNS によって、人々は公的な事柄について自由に発言し、議論することができるようになりました。ネガティブな意見も含めてですが、それも「文脈からの解放」という断片化の成果です。

もしインターネットが存在しなければ、とりわけ日本では、多くの人々は政治的な発言などできなかったでしょう。現実の社会では、自分の会社や労働組合、地域社会といった所属コミュニティの方針に従うのが普通でした。そこから抜けられない限り、自分自身の政治的意見を表明するのは難しかったのです。インターネットは、そうした社会的文脈から個人を切り離す力を持ってきましたが、断片化が行き過ぎれば、移ろいやすいその場しのぎの意見に流されたり、情報操作に振り回されたりするという危険もあります。

重要なのは、「可変性」と「安定性」のバランスを取ることです。AI は使い方次第で、社会の分断に一定の歯止めをかけ、一貫性や繋がりを回復する役割を果たせるかもしれません。民主主義の原理に立ち返って考えると、民主主義が元々抱えてきた「刹那的で断片化しやすい性質」、そしてデジタル技術がそれを加速させてきた傾向を、逆に AI が補い、繋ぎ直す可能性もあるのではないかと——。私はそこに期待すべきだと思えます。

## 「責任ある AI」をどう仕組み化するか

**菊池** AI ガバナンスについて伺います。インターネットは「自律・分散・協調」という理念の下、数十年かけて発展・成長を遂げました。AI についてはどうでしょうか。AI はインターネットのあり方さえ変貌させてしまうでしょうか。

**穴戸** インターネットを形成するそれぞれのネットワークは、自分の中で起こる問題を自分で処理する「責任」を持ち、その上で、他のネットワークとつながるときには「協調」して情報をやり取りしてきました。この仕組みの裏側には常に「責任」と「信頼」がありました。各ネットワークがきちんと責任を果たすからこそ、全体として分散型の構造が効率的に機能し、国家や企業のような中央集権的な仕組みに対してインターネットの優位性が保たれてきたのです。

しかし、インターネットが拡大する中で、この「責任」の部分が揺らいできました。違法な海賊版サイトの横行問題などはその一例です。そこに AI が登場し、今後は AI がこの責任のあり方をどう変えるかが議論の焦点になると思います。

例えば、あるネットワークに「悪質な AI」がつながっていて、その AI が大量の有害な情報を発信しているのに、そのネットワークが接続を遮断しなければ、「責任を果たしていない」と見なされ、そのネットワーク自体がインターネットから切り離されるかもしれません。こうした事態は、国ごとにインターネットを分断する「スプリンターネット」（サイバー空間への自由なアクセスに対し、国家間や対立する地域間で何らかの制限を加えること）を正当化する口実にもなり得ます。

一方で、逆の可能性もあります。AI を使うことで、ネットワークは今まで以上に責任を果たしやすくなるかもしれません。既にデータのやり取りの処理には AI が活用されていますし、AI がネットワークの効率や安全性を高めることもできます。「責任ある AI の利用」が進めば、「自律・分散・協調」の仕組みそのものを強化することも可能なのです。

つまり、インターネットの自律・分散・協調の基盤には「責任」があり、AI はその責任を壊す方向にも、支える方向にも働き得ます。だからこそ、AI ガバナンス、すなわち責任ある AI をどう仕組み化していくかが、今後もインターネットが「信頼」できるかどうかを左右することになると思います。

**菊池** インターネットの「自律・分散・協調」を、フランス革命時に唱えられた「自由・平等・博愛」と比べると、インターネットには「平等」という要素が弱かったのではないかと感じます。

**穴戸** 「自由・平等・博愛」というスローガンが登場した歴史的背景を考えると、アンシャン・レジーム（古い体制）、つまり身分制社会に対する抵抗のスローガンでした。そのスローガンには、特に「平等」という概念が強く織り込まれていました。身分制を否定し、自由で平等な個人から構成される新しい社会を築こうという強い意味を持っていたのです。

インターネットにおける「分散」の理念も、同じように既存の集権的な仕組みに対するアンチテーゼでした。近代以降の情報処理技術の進展は、国家や企業の集権化、あるいは独占性を強めてきました。大型コンピュータによる集中電算処理が典型です。そうした流れに対抗して、インターネットは「分散」という価値を掲げました。これは、近代社会がアンシャン・レジームに対し、「平等」という価値を掲げて戦ったのと同じ構造であると思います。

ただし現実には、フランス革命後にブルジョワジーの経済的独占が生まれてしまったのと同じように、イ

インターネットも「分散」を掲げながら、GAFAのような巨大プラットフォーマーによる集権化を生んでしまいました。その結果として、また「分散を取り戻さなければならない」という議論が出てきています。ただし、「平等」という言葉の意味が時代とともに変わってきたのと同じように、デジタル資本主義における「分散」も再定義されつつあります。たとえばDAO（分散型自律組織：ブロックチェーン技術を活用して自律的に運営される組織形態）のような仕組みがその一つです。これは近代半ばに平等の意味が問い直されたのと同型の議論をしていると感じます。

## AIに期待したい、「部分最適」と「全体最適」の両立

**菊池** 5年先を見通すことも難しいほど変化のスピードが加速していますが、宍戸先生のご専門である憲法は、5年や10年といったテクノロジーの変化や人口・社会の動態を超えた、安定性の源泉だと思います。変わらない理念や価値観を守る憲法と、変化する現実との間を橋渡しするのが「情報法」などの役割であると考えて良いでしょうか。

**宍戸** その通りだと思います。憲法は社会や技術の変化に関わらず、人間の尊厳や民主主義、法の支配、平等など、容易に変えるべきでない価値を体現しています。同時に、憲法は生きた存在、いわゆる「Living Constitution」でもあり、時代の変化とともに進化すべきものであるとも考えられています。アメリカの連邦裁判所の判例のように、実務を通じて憲法の中身を更新することもあります。

EUでは加盟国全体の憲法に当たる欧州連合基本権憲章を定めた上で、GDPR（一般データ保護規則）やDSA（デジタルサービス法）、DMA（デジタル市場法）、AI法などが整備され、デジタル立憲主義の流れも生まれています。日本においても、憲法が体現する価値を実現するために、デジタル技術を使って民主主義を安定・高度化させる余地は大きいと思います。人口減少や資源制約の中で、行政の効率化を法の支配の下で進めることも可能です。

ただし近代憲法の前提は「人間が個人であること」です。AGI（汎用人工知能）やASI（人工超知能）の登場によって、人間の存在そのものが揺さぶられる可能性があるとするれば、憲法だけの問題ではなく、近代社会そのもののあり方が問われることになります。そうした変化を法で引き受けて議論する場合、「人間個人という存在」を前提とする現行の憲法がそのまま通用するのか、再検討が必要になるでしょう。

実際、ニューロサイエンスの発達やエンハンスメント技術によって、人間が機械と一体化する可能性が見えてきています。人間の脳も機械とつながってくるかもしれません。そうなると、人間の尊厳という観念が根本から脅かされるかもしれません。これまでは倫理や法によって一定の抑制ができましたが、AIに対してそれが通用するのか。ASIレベルになると、人間の裏をかいたり、欺いたりすることも簡単にできるよう。従来の方法では対応できなくなるかもしれません。

**菊池** 国民国家はAIの登場によってどうなっていくのでしょうか。

**宍戸** 変わると思います。物理的に鎖国すれば国民国家を維持できるかもしれませんが、言語やコミュニケーションのレベルでは国境の壁が下がりつつあります。その意味で、国民国家の成立に関わる言語やコミュニケーションのあり方が変わってくるでしょう。

一方で、グローバル化やデジタル化に対する反発もあります。これは社会全体や個人の心の中にも存在し、私たちはそれとどう折り合いをつけるかを考えざるを得ません。これは、文明開化の時と同じような変化のプロセスだと思います。

私が期待しているのは、部分最適と全体最適を AI が両立してくれることです。人間は自分に都合のよい選択と全体の規律のどちらかを選ぶことしかできませんでしたが、AI ならそれを同時に実現できる可能性があるのではないかと考えています。

**菊池** 合成の誤謬を、AI が回避するというのでしょうか。

**穴戸** そうなることを目指していくべきです。形式的で画一的な処理は周辺化を生むだけですが、AI を使えばそうならず、部分最適を保ちながら全体最適を実現できるかもしれない。だから、自律・分散・協調の「分散でありオープンである」という考え方が大事で、AI 時代においても目指すべき方向だと思います。

**菊池** たいへん貴重なお話をいただき、まことにありがとうございました。

## 穴戸 常寿 / George Shishido

東京大学大学院法学政治学研究科 教授

1997 年東京大学法学部卒業。東京大学大学院法学政治学研究科助手、東京都立大学法学部助教授、一橋大学大学院法学研究科准教授等を経て、2013 年より現職。専門は憲法・情報法。主要著作として『デジタル・デモクラシーがやってくる!』（共著、2020 年）、『AI と社会と法』（共編著、2020 年）、『憲法学読本（第 4 版）』（共著、2024 年）等。現在、東京大学大学院法学政治学研究科附属法・政治デザインセンター長、国立情報学研究所客員教授、個人情報保護委員会委員、衆議院議員選挙区画定審議会議員等。

インタビュー実施：2025 年 7 月 23 日

# 創作エコシステムが崩壊の危機に

上沼 紫野

弁護士

菊池 AIが文章を書き、アニメ作品風のイラストを易々と生成し、楽曲の作詞・作曲までこなせるようになり、人間の特権だった「表現」や「創作」の領域が侵食されているように思います。これまでに構築されてきた著作権、知的財産権などの法制度、言い換えれば、表現や創作のエコシステムに、AIはどのような影響を及ぼすのでしょうか。

## AIによる「6割代替」でクリエイター・エコシステムが崩壊へ

上沼 著作権というものは、英語で copyright というように出版権（印刷による複製権）の保護・制限のために確立されてきたものですが、デジタル時代に入ってかなり無理をしてきたと思います。昔は複製することに多大な労力がかかっていたので仮に著作物として保護されなかったとしても、そうそう簡単に複製することはできませんでした。ところがデジタルでは複製コストがほとんどかかりません。複製しようと思えばいくらでもできてしまう。そのため、「これは著作物か否か」という線引きが非常に大きな意味を持つようになったのです。ところが、その基準は明確でないまま、ごまかしながら凌いできたのです。

そこにAIが登場し、「誰が創作者なのか？」という問い、つまり、人間が創ったのか、AIが作ったのか、新たに浮かび上がってきました。「そもそも、何を保護するのか？」という原点に立ち返らなければならない状況にあると感じています。

もう一つの大きな問題は、人間のクリエイターが要らなくなり、育てられなくなり、いなくなってしまうのではないかということです。創作物に限らない多くの仕事において、AIが「6割」くらいまでをこなせるようになってきていて、それはすなわち、これまでその6割の部分を担当してきた人間がいらなくなるということです。しかし、どんな仕事でも下済みの6割のところまで仕事を覚え、慣れ、腕を磨いたうえで、その先に挑めるようになっていったわけです。ところが、その6割がAIに代替されてしまうと、若手が経験を積む機会がなくなり、成長の道筋自体が失われてしまう。

しかも、AIは人間より成長スピードが速い。人は寝たり休んだりするけれど、AIは止まらずに学び続けられる。だから同じスタート地点に立ったとしても、人間はAIにとっても追いつくことができません。クリエイターを目指してその道に入っても、初心者レベルのことはAIが全部やってしまうので、創作で生活を成り立たせることが難しくなってしまうのです。AIを道具として使いこなし、AIが生成できるものを超える新しいものを



創り出せればクリエイターとして生き残れるかもしれませんが、そこに至るエコシステムが崩壊の危機に直面しているように思います。著作権の保護以前の問題です。

**菊池** 「複製（リ・プロダクション）」ではなく、「創作（プロダクション）」そのものが問題であり、何を保護するかという前提が変わってきているわけですね。

**上沼** そう思います。私も、5年くらい前まではまだ楽観的で、「一種の産業革命だ」と考えていました。その頃懸念していたのは「データ資本主義」の加速でした。財力のある人ほど質の高いデータを集めてAIを強化できるので、AIによる寡占化が進むだろうと。それでも「データ成金」みたいな人が人間の手による古き良きものにお金を投じてくれば人による創作が残るかなと思っていました。

しかし、今の状況を見るとルネサンス時代のようなことにはなりそうもありません。AIで事足りるのにあえて人間がつくるものに投資して保護することをステータスとして尊ぶ感覚が生まれてくるかどうか……。そこに頼るのは、難しいかもしれません。

例えばアニメ制作の現場などは元々賃金が必ずしも高くなかったのに、AIがその仕事を根こそぎ奪ってしまうかもしれない。アニメーターという仕事が職業として成り立たなくなってしまうかもしれません。そうすると、創作者、表現者のすそ野がどんどん狭くなり、新しい才能を芽生えさせる土壌が失われていく。そうした負のスパイラルに入っていくことを一番懸念しています。

## 多様化から単一化・画一化への逆行のおそれ

**菊池** AIの出現によって、エンタテインメントや芸術はどうなっていくのでしょうか。

**上沼** エンタメや芸術は基本的に主観の領域にあるので、多種多様であることが是であり、これまではそうした方向に進んできました。しかし、AIの出現によって全体的に見ると「単一化」「画一化」の方向に逆行してしまうのではないか。AIはデータに基づいて結果を出すので、みんなが良いと思うようなものが繰り返し再生産されていきます。行き着く先は、多様性の拡大よりも凝縮された単一化なのではないか。

しかも、AIは米中欧のグローバル企業が強いので、文化的な価値観がそうした国々の人たちが好む方向に偏ってしまうかもしれません。個人的には、それが一番不安です。

結局はお金が必要なんです。日本のサブカルチャーは熱心な少数のファンがお金を出すことによって成立してきました。一人のファンが10枚、20枚のCDやチケットを買ったりする「熱量」によって支えられている。でもデータ資本主義においては、一人の「熱量」よりも、総額でいくら「課金」できるかが重要なのです。少数の熱狂的ファンが支える仕組みはデータ資本主義の時代には意味を持ちにくい。

目に見える形で市場原理が支配してしまうのは、正直怖いですね。AI時代のデータ資本主義によってエンタメと芸術のコモディティ化が進んだとして、人間はそれに順応してしまうのか、それとも、反発してカウンターを当てにいくのか、そもそもAIに対してカウンターを当てられるのか……。あまり楽観的にはなれないというのが正直なところでは。

## 外国製 AI を日本の文化・社会が受容できるか？

**菊池** AIは「コミュニケーション」をどう変えるのか、家族・友達になるのか、それとも武器なのか、需要や地域の違いによって受容性が変わってくるのか、日本の社会にとってはプラスなのかマイナスなのか。

**上沼** 横浜商科大学の田中辰雄先生が統計をだされていました<sup>1</sup>が、日本人は他国に比べて AI やロボットへの親和性が高いと思います。背景には八百万の神という考え方がありますし、鉄腕アトムのようなアニメの影響もあるかもしれません。AIBO が登場した時、ロボットをペットにするという発想に世界中の人たちが驚いたのですが、日本人にとってはそれほど違和感もなく自然に受け入れられました。

ただし、親しみの情が濃い一方で、恐れ of 気持ちも強いことも同じ統計で出ています。森羅万象の一部として親しく感じると同時に人間とは違うもの、もしかしたら超越するものへの畏れというものにも敏感なんだと思います。西洋的な価値観ではロボットや AI は「道具」という認識が強く、人間と仲良く、親しくするという発想は薄いように感じます。文化的、社会的な違いが顕在化してくるのではないかと思います。

**菊池** AI に対するカウンターとしての「人間ファースト」という流れも出てくるかも。

**上沼** あり得ると思います。厳しいのは、日本企業が AI で強みを持っていない点です。もし「ここは国産 AI が絶対に強い」という領域があれば安心感も違ったでしょうが、残念ながら、現状はそうではありません。だからこそ外国製 AI への恐怖があるのだと思います。市場規模が小さく人口も減っていく日本は、相手にされない可能性がある。

日本的な曖昧さやグレーゾーンを許容するような AI が全く無いとすると、八百万の神のような感覚で付き合えないでしょうし、文化的なハードルが生まれるかもしれません。

**菊池** 明治の文明開化では「これをやってはいけない」「これを変えよ」と外からルールが押し付けられましたが、あの時代は日本には活力があった。

**上沼** 本当にそうですね。明治の人たちは。欧米諸国に追いつこうと懸命に頑張りました。現代に生きる私たちは大いに見習いたいところです。AI への需要はその国の活力と密接に関わっているように思います。高齢化する日本では AI を駆使した自動運転などへの切実なニーズもあります。AI をどのように、どれだけ受容できるかは、それぞれの国・地域における社会状況と直結していると思います。

## 自分の考えは本当に自分のものなのかを問い続けよ

**菊池** 5年後、AI と表現・創作はどうなっているのでしょうか。5年前、上沼さんは比較的楽観的に見ていたとおっしゃっていましたが……。

**上沼** 実は5年前も本当は悲観的だったんです。ただ、悲観してばかりではいけないと思ってあえて明るい側面を見ようとしていたというのが正確なところですよ（笑）。

この5年で大きな変化が起きました。これからの5年ははるかに大きく変化することは確かでしょう。だからこそ、私たち AI ユーザーが備えるべきは、「自分の考えが本当に自分のものなのかを常に問い続けること」だと思います。入ってくる情報が偏れば、考え方も自然に偏って自分自身が気づかないうちに認知が歪められてしまう。そこへの警戒は絶対に必要です。

**菊池** パーソナル AI エージェントは、フィルターバブルやエコーチェンバーの究極の形かもしれません。

**上沼** AI エージェント自体をどう信頼すればよいのか、まだ誰も分かっていません。そして結局はお金がある人ほど優秀なエージェントを使えるようになるでしょう。それは決して明るい未来とは言えません。

---

1 「日本はなぜ AI に好意的なのか」、田中辰雄（横浜商科大学 教授／国際大学 GLOCOM 主幹研究員） <https://www.glocom.ac.jp/publicity/discussion/9139>

## AI 資本主義の大波に日本は飲み込まれるだけか？

**菊池** インターネットは「自律・分散・協調」と「マルチステークホルダー」という精神に基づいて成長してきましたが、この精神は AI 時代に受け継がれるでしょうか。

**上沼** 引き継がれるかもしれませんが、それをそのまま実践した結果が、今のデータ資本主義だとも思っています。つまり、精神は残ったとしても、その行き着く先が本当に望ましい形かどうかは分からないということです。どうしても商業的に利益の出る方向へ流れざるを得ないでしょう。規制がなければなおさらです。結果として「AI 資本主義」に収束してしまうのではないかと危惧しています。

**菊池** AI 資本主義の暴走を抑えるために、法律はどの程度有効でしょうか？

**上沼** 正直なところ、国レベルの法律では効力が限定的だと思います。日本が何を言っても、世界全体から見れば影響力はとても小さい。国際的な足並みが揃わないと間に合わないでしょうね。条約レベルでの対応が必要でしょう。国ごとのバラバラな法律よりはましだと思います。EU が地域単位で規制を整えようとしているのは、その一例です。日本よりも積極的に動いていると感じます。

**菊池** 国際連合には期待できないのでしょうか。

**上沼** 国連は第二次世界大戦の戦後体制を前提に作られており、その構造を固定化してしまいました。だからウクライナの問題にしても身動きが取れない。今の国連だと、あまり期待できないのでは？というのが正直な印象です。

**菊池** 国内法を武器にしてきた弁護士である上沼さんが、条約を重視するというのは皮肉に聞こえました。

**上沼** 確かに（笑）。ただ、日本は国同士の法律競争になると弱いのです。法律そのものの問題というより、それは国力の問題でもあるからです。人口が減り、国力が落ち、言語的にも文化的にもマイナーな国が、AI 時代に存在感を示すことができるのか、大きな流れに飲み込まれてしまうのではないか——。そのリスクにはもっと危機感を持つべきだと思います。

**菊池** 貴重なお話をいただき、まことにありがとうございました。

### 上沼 紫野 / Shino Uenuma

LM虎ノ門南法律事務所 弁護士

LM虎ノ門南法律事務所。千葉県立長生高校理数科、東京大学法学部卒。Washington University in St.LouisでLLM取得。弁護士、ニューヨーク州弁護士。米国Perkins Coie法律事務所、FTCでインターンの後、知的財産、IT関連、国際契約等の業務を主に行う。最高裁判所司法研修所刑事弁護教官（2012－2015）、子ども家庭庁青少年インターネット環境の整備等に関する検討会委員（2015－2025）、サイバーセキュリティ戦略本部サイバーセキュリティ推進専門家会議（2025）、情報セキュリティ大学院大学客員教授、一般社団法人安心ネットづくり促進協議会理事等を務める。

インタビュー実施：2025年7月22日

# AI ガバナンス、国際舞台でのかけ引き

飯田 陽一

総務省 参与

**菊池** 飯田さんは総務省で、AI ガバナンスに関する国際的な調整業務に長く携わってこられました。AI に関する国際的合意形成のこれまでの経緯と今後の展望についてお聞きかせいただけますか。

## 2016 年 G7 伊勢志摩サミット：日本がリードした議論

**飯田** AI に関する国際的なルール作りは、日本が最初に提案したことなのです。2016 年の G7 伊勢志摩サミットの直前に開催された情報通信大臣会合に先立つ WG で私は議長を務めていて、「AI は社会に大きな影響を及ぼすから、国際的なルール作りを始めるべきだ」と日本から正式に提案したのです。

ところが反応は、ほぼ総スカン（苦笑）。「インターネットや通信と違って、AI はコンピュータ技術の一つの分野に過ぎない。情報通信大臣会合で扱うテーマではない」とバツサリでした。当時はまだ、AI を世界の問題として取り上げる意識が高まっていなかったのです。

しかし、日本の立場は違っていました。2016 年早々、「AI ネットワーク化検討会議<sup>1</sup>」を総務省が立ち上げ、かなり先取りした議論を始めていました。経済産業省も同様の問題意識でした。そうした先行的議論の中から、AI を巡る検討は日本だけでやっていたはならず、国際的に広げていく必要があるという問題意識が生まれていたのです。

G7 情報通信大臣会合では、当時の高市早苗総務大臣が、「AI は国境を越えて社会に大きな影響を及ぼす技術であり、情報通信に関わるテーマなので、来年以降も議論を続けますよね?」と押し迫り、議長国がそこまで言うのを各国も嫌だとは言えず継続的に議論することが、口約束レベルではあったにせよ、なんとか決まったのです。あの場で日本が踏ん張らなければ、AI に関する国際的議論は数年遅れていたかもしれません。

**菊池** その後、どのように進展していったのでしょうか。

**飯田** 2017 年の G7 議長国はイタリアでした。日本からイタリアの事務方には「AI はとにかく大事。昨年の約束もあるので、議題としてしっかり取り上げてほしい」と懸命に働きかけました。最初、イタリア側は「本当にやるのか?」と必ずしも積極的ではなかったのですが、自動運転や医療など、AI が社会に与える影響は非常に大きいという認識と理解が徐々に広がっていきました。その頃から「人間中心の AI (human-centered AI)」という言葉が使われるようになり、倫理や透明性をどう確保するかという議論が国際的な場でもされるようになっていったのです。

そして 2018 年、カナダで開催された G7 シャルルボア・サミットで「AI の未来のためのシャルルボワ・共通ビジョン (G7 AI ビジョン)<sup>2</sup>」の合意に至り、「人間中心の AI」という概念が明確に合意されました。

1 [https://www.soumu.go.jp/main\\_sosiki/kenkyu/iict/index.html](https://www.soumu.go.jp/main_sosiki/kenkyu/iict/index.html)

2 [https://www.mofa.go.jp/mofaj/ecm/ec/page4\\_004124.html](https://www.mofa.go.jp/mofaj/ecm/ec/page4_004124.html)

国際的な合意に基づく AI 原則が初めて文書化されたのです。

その舞台裏でも攻防がありました。AI に関して守るべき原則を列挙する中に「データ」という項目があったのに対して、ある国の外交官が「我々は人間の尊厳と AI の関係を議論しているのであって、データ云々とは関係ない」と異議を唱えたのです。危うく「データ」に関する記述がごっそり削られかけて、もうびっくりしまして、日本の外務省に頼んでなんとか押し戻してもらいました。AI とデータは密接不可分という考えは、当時はまだ議論をしていた外交官の間でも共通認識になっていなかったのです。AI の健全な発展と開発にはデータが果たす役割が大きいという観点がかろうじて残ったことで、そこから徐々に理解が深まってきたという経緯があります。

## 2019 年 OECD 「AI 原則」のベースは日本モデル

菊池 G7 以外の国際機関では、どのような状況だったのでしょうか。

飯田 OECD（経済開発協力機構）でも最初は同じような反応でした。G7 での議論を受けて 2016 年に日本が AI の議論を提案した時、OECD 科学技術イノベーション局長だったアンドリュー・ワイコフ氏は、インターネットならともかく AI を積極的に扱うスタンスではなかった。そこで、総務省から OECD に人を送り込んで AI 専門タスクフォースの立ち上げに持ち込み、先述したネットワーク化社会検討会でまとめた「AI 開発原則」を持ち込んで議論の土台にしてもらいました。この AI 原則の日本モデルには、「人間中心」「透明性」「安全性」などが盛り込まれており、後の国際原則を先取りする内容でした。

そして、2019 年に OECD の「AI 原則（AI に関する理事会勧告）」が採択されました。日本がかなり汗をかいたこともあって、その内容には日本の AI 開発原則の要素が色濃く残りました。事実上、日本がそのベースを提供したものだと言っていると思います。

2020 年には、「GPAI（Global Partnership on Artificial Intelligence、AI に関するグローバル・パートナーシップ）」が発足します。これは、OECD の AI 原則に基づきマルチステークホルダーによるプロジェクトによって開発・利活用を推進するための国際的な官民連携組織です。提唱国であるカナダ、フランスに続き、2023 年には日本にも GPAI 東京専門家支援センター<sup>3</sup>が設置されています。ここでも日本の立場は一貫していました。すなわち「規制は必要最小限にとどめる」「できるだけ自由な開発・利用を確保する」という考え方です。政府が上から規制をかけるのではなく、民間や研究コミュニティが自主的にガイドラインを作り、実践する形を目指したのです。

2020 年から本格的にプロジェクトが走り出し、例えば「責任ある AI の実装」「データガバナンス」「イノベーションと商用化」といったテーマごとに作業部会が設けられ、産学の専門家が共同研究を始めました。当時の雰囲気はまだ楽観的でした。AI の負の側面よりも「どう利活用を進めるか」に関心が強かった。特に米国は GAFA など大手プラットフォーマーを抱えていますから、規制強化には当然慎重



3 <https://www2.nict.go.jp/gpai-tokyo-esc/>

でしたし、日本も「官がトップダウンで縛るとスタートアップや研究現場のスピード感を殺してしまう」という危機感を共有していました。

日本や米国、英国は「規制は必要最小限に」「自主的なガイドラインで十分対応できる」という認識を共有していましたから、この「ソフトロー」的アプローチが国際的な標準になり得ると期待していたのです。

## EUの「AI規制法」でG7内分裂の危機

ところが、このソフトロー的アプローチの流れに水を差したのがEU（欧州連合）です。ふと気づけば、いわゆる「AI法」の検討を始めていたのです。非拘束、非規制、できるだけ自主規制に任せるという方向性で一緒に取り組んできたと思っていたので、正直驚きました。

担当者に問うと、「EU市民はAIの負の側面を恐れている。しっかりした規制がないと利活用も進まないという考え方もあるので、どちらが良いか検討しているのだ」と答えるので、それからは事あるごとに「過剰規制はEUのためにも世界のためにもならないから絶対にダメ。AIの発展の速さや利活用の多様性を見ればとてもじゃないが水平規制は無理。気をつけたほうがいい」と、失礼だとは思いつつもかなり率直に言い続けました。EUの担当者は「分っている。過剰な規制はしない。必要な規制しかしない」と答えていましたが、のれんに腕押しのような感じで・・・その裏ではAI法制定に向けた検討がどんどん進んでいたのです。

非常に危惧したのは、EUのAI法が過剰規制に走って、日米の非拘束型ソフトローのフレームワークと分裂してしまうことでした。G7の外側には中国をはじめとするプレイヤーがいてまた違うことをやっていることを考えるとG7内がバラバラになってしまうのは非常にまずいと思いました。そこで、2022年の段階で、日本が議長国となる2023年のG7でAIをもう一度取り上げることにしたのです。

## 2023年：ChatGPTの衝撃が生み出した「広島AIプロセス」

G7群馬高崎デジタル・技術大臣会合<sup>4</sup>に向けて、各国の事務方で準備を進めました。深層学習や機械学習のAIをどのように管理し、開発促進し、利活用するかについて、いかにグローバルな枠組を作っていくか議論を重ねたのですが、特にEUに自由なイノベーションが大切であるという点を改めて認めさせようとしていました。

ところが、何回目かの事務方会合の場で合意文書のドラフトを練っている時、EUの担当者から会議室の外に呼び出されたのです。曰く、「EUとしてはこのドラフト内容は飲めない。EU内部でAI法を検討している中で、“法規制よりもイノベーションが大切”というような文言が残ったら自分はブラッセル（EU本部）に帰れない。絶対に受け入れられない」と。

そこから、どこまでなら折り合えるかの探り合いが始まりました。まずは、OECDのAI原則、自由民主主義、基本的人権の保護といった共通の価値観を尊重したガバナンスの枠組みが必要であるということを確認しました。その理念を実現するためのアプローチが国・地域ごとの歴史的・文化的・経済的背景や状況によって違ってくるのは許容するが、G7各国が分断してしまうことは回避する。そのためには、

4 [https://www.soumu.go.jp/joho\\_kokusai/g7digital-tech-2023/](https://www.soumu.go.jp/joho_kokusai/g7digital-tech-2023/)

透明性を確保し、各国フレームワークの相互運用性を維持していくべきである——。そういうスタンスに切り替えたのです。

EU が離脱してしまわないよう気を使いながら議論を進めていた時のことです。2022年11月に ChatGPT3.5 をリリースしていた OpenAI が、翌年3月には ChatGPT4 を正式リリースし、さらにサム・アルトマン CEO が4月に突如来日して岸田文雄総理（当時）と面会。新聞の一面を毎日 AI の記事が飾りました。突然、「生成 AI」を抜きにして AI ガバナンスの議論は成立しなくなりました。

ところが、生成 AI をそれまでの AI と一緒にして論じられるのか、生成 AI 特有のリスクやポテンシャルは具体的に何なのかなど、詳しいことはその時点では分かりません。デジタル大臣会合は2023年4月末に迫っています。それまでに結論を出すのはとても無理でした。

そこで、大臣会合後も継続して議論しようということになったのですが、驚いたことに一番前のめりだったのは腰が引けていたはずの EU だったのです。EU 代表は、「すぐに G7 で議論を始めるべきだ。日本の議長下でぜひとも結論を出してほしい」とまで言うのです。生成 AI の強烈な衝撃のおかげで空気がガラリと変わりました。2016年の G7 伊勢志摩サミットの冷ややかな空気とは真逆です。私自身、体の中で血が逆流するような、そんな気持ちの高まりを感じた瞬間でした。

そんな経緯でサミット後も関係閣僚で議論を継続することになったのですが、広島サミットで、ならば名前をつけようということになりました。それが「広島 AI プロセス<sup>5</sup>」だったのです。

## 2025年：バンス演説で「AIの安全性」の議論が停滞

**菊池** 広島 AI プロセスの背景に AI ガバナンスの分断危機があったとは……。生成 AI の登場が逆に救いになっていたということにも驚きました。広島 AI プロセスの成立から現在まで、国際的コンセンサス形成はどのように進んだのでしょうか。

**飯田** 進んだ面もありますが、現在は停滞していると言ったほうがいいかもしれません。

2024年8月に「EU AI Act」、別名 AI 規制法が施行されました。ただし、EU 加盟国は必ずしも一枚岩ではなく、特にフランス政府は強すぎる規制に懐疑的だと言われています。実際、2025年2月にパリで開催された「AI アクションサミット<sup>6</sup>」はフランスとインドが共同議長を務め、AI 開発における国際協力を推進すること、過剰な規制には反対する姿勢であることをアピールしました。

ところが、米国からのゲストとして招かれていた J・D・バンス副大統領はさらに踏み込んでこう切り出しました。「私は AI の安全性 (safety) について話すためにここに来たのではない。AI の機会 (opportunity) について話しに来たのだ」。この言葉は EU の AI 法を批判・牽制するものでしたが、それにとどまらず AI の安全性に対する取り組み、AI ガバナンスそのものを否定するように捉えられました。

このバンス演説以降、今まで G7 や OECD で作ってきた AI の「安全性」に関するフレームワークは協調的に運営できなくなるのではないかという危惧が広がり、明らかに議論が停滞してしまいました。さらに、様々な政策文書から「AI の安全性」という言葉が取り除かれたり、「信頼性 (reliability)」といった別の言葉に置き換えられたりして、まるで言葉狩りのような様相を見せるほどでした。

5 <https://www.soumu.go.jp/hiroshimaaiprocess/>

6 <https://jp.diplomatie.gouv.fr/ja>

そして、2025年7月に「AMERICA'S AI ACTION PLAN<sup>7</sup>」が発表されました。バイデン前政権の安全規制の撤廃、米国製 AI の輸出促進、データセンター建設の許可手続き迅速化など、徹底的なイノベーション促進策となっています。そして、米中対決という大きな構図の中で、どちらの側につくのかという選択を他国に迫る内容になっているように見えます。

## 広島 AI プロセス・フレンズグループに「G7」が付かない理由

**菊池** AI ガバナンスを巡る国際的議論は、今後どのように進んでいくのか、いくべきなのか——。飯田さんはどうお考えですか。

**飯田** 非常に難しい質問です。理想を言えば、米中が、相容れないところはあったとしても、ちゃんと話ができ、ある程度の相互理解ができれば良いなとは思いますが、簡単なことではないでしょう。

G7 では、広島 AI プロセスを含めて、「G7 のためのルールではなく、グローバルなルールを目指していこう」という意識をもって私たちは頑張っていました。ですから、「広島 AI プロセス」には「G7」という文字が入っていないのです。G7 以外の国々から嫌われたら仲間の輪を広げられないので、あえて入れなかったのです。

広島 AI プロセスに賛同する国を増やしていくために、「フレンズグループ」という枠組みを 2024 年 5 月に日本主導で立ち上げました。ところが、準備段階でいざどの国を入れるのかという話になったとき、各国の反応が渋いのです。シンガポール、韓国、オーストラリアくらいまでは大丈夫なのですが、その先になるとなかなか賛同が得られない。何が心配かという点、一つは G7 レベルの議論についてこられないような国が入ってくるとカオス状態に陥って議論が進まなくなるのではないかという懸念。もう一つは、自由民主主義的な価値観と相容れない国々を安易に入れてしまうと根本的に議論がかみ合わなくなり、枠組みが崩壊してしまうのではないかという懸念です。

そうした心配もわかります。しかし、中国はものすごい勢いで仲間づくりを進めているのに、そんなことを言って止まっていたら勢力図が書き換えられてしまうかもしれません。それに、私自身は官邸から「やれ」と言われているので立場上やらないわけにはいきません（笑）。最後は、「ならば日本が勝手にやる」と啖呵を切って始めたのがフレンズグループなのです。G7 各国も付き合いで入ってはいますが、日本がいれば永年議長になってどんどん声をかけていくことにしました。現在では、発展途上国を含めて 61 개국 + EU が加盟してくれています（2026 年 2 月時点）。

国際連合という場もあるではないかとおっしゃるかもしれませんが、中国が国連加盟国を本気で先導すれば、G7 は数で負けてしまうでしょう。中国と途上国にとって都合の良い国際ルールが出来上がってしまい G7 が作ったものは目もかけられないということになりかねません。特に政府がトップダウンで規制を課し、人権やプライバシー、基本的な自由を尊重しない制度が世界に広まる。それは避けなければならない。ですから、国連の場で AI が議論された時、日本や G7 の主張に耳を傾け理解してくれる国を少しでも増やしておきたいのです。フレンズグループには何の拘束力もありませんが、普段から基本的な考え方を共有しておけば、いざという時に話くらいは聞いてくれるのではないかと願って、ここまで頑張ってきました。

7 <https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf>

## 理想はマルチステークホルダーの IGF 型

**菊池** G7 でも OECD でも国連でもないとなると、やはり IGF（Internet Governance Forum、主催は国連<sup>8</sup>）のようなマルチステークホルダー方式によるコンセンサス形成を目指すべきなのでしょう。

**飯田** 政府だけでなく、全ての関係者が関与して、意見を言って、ガバナンスが構築・運営されていくのが理想的だと思います。日本ではこの9月1日に AI 法（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）<sup>9</sup> が全面施行されましたが、これは EU の AI 規制法とは違い、しっかり運用すればマルチステークホルダーによるガバナンス、共同規制的な方向付けが可能であり、一つの理想形だと思っています。

ただし、これをそのままグローバルに展開できるかという点、かなり難しいと思います。民主主義体制の国々となら一緒にやれるかもしれません。しかし、政府が圧倒的な力を占めている体制（権威主義国家）と接合するためには、政府にある種のタガをはめるような強制力ある規律をかける必要があるかもしれません。それが受け入れられなければ一体化は難しいでしょう。言い換えると、法の支配、個人の自由、プライバシーといった守るべき絶対的価値がある国とは政府はどこまでやっていいかという話し合いが成立しますが、それが無い国とはゼロからの議論になってしまうのです。ですから、G7 や OECD の延長線上に絶対解はないと思います。国連で統一的な AI ガバナンスのフレームワークを作ることも難しいでしょう。

やはり IGF のようなマルチステークホルダーが集まって議論する場が必要だと思います。AI の課題を自由に議論して、それを持ち帰って各国・地域で制度や枠組みを整える。それをまたマルチステークホルダー間で共有してということを繰り返しながら、グローバルな枠組みに向けて一緒に議論を続けるというイメージです。IGF には中国の関係者も参加していますからね。

2年後くらいに目指しているのは、IGF の AI 版ができていてマルチステークホルダーによる AI ガバナンスの議論が始まっている状態です。そして、最後は IGF と統合していくのが自然だと思います。最初から IGF でやるという考え方もあると思います。

## アメリカが大きな不確定要因に

**菊池** IGF の AI 版、“AIGF” をつくるというアイデアはとても興味深いですね。中国を巻き込むということも重要ですが、かつては世界のリーダーだったけれども今では“ならず者”のようになってしまったアメリカをフラットな関係性で引き込むための仕組みが必要になっているのだなと感じました。

**飯田** 確かに、アメリカが不確定要因になってしまっている現状はあります。個人的には、AI などデジタル分野では、表面上は威勢のいいことを言っても実際にはそれほど極端なことはしないだろうと思っているのですが、予測不能性は高まっていて、事あるごとに手探りで向き合わなければなりません。広島 AI プロセスの交渉も「日本が米欧の間に入ってうまく調整した」と言っていますが、本当は日本がアメリカの力を借りて EU を引き込んだというのが実態です。

ですから、アメリカは表面的にはヤンチャなポーズを取り続けているけれども水面下では今までと同じように日米で協力していける、という一点さえ確認できれば、日本はアメリカの“フルスタック”という言葉

8 [https://www.soumu.go.jp/menu\\_seisaku/ictseisaku/igf/index.html](https://www.soumu.go.jp/menu_seisaku/ictseisaku/igf/index.html)

9 [https://www.cao.go.jp/press/new\\_wave/20251003.html](https://www.cao.go.jp/press/new_wave/20251003.html)

に乗っかって日米混成 AI をどんどん展開できる。これはむしろ大きなチャンスになると思います。

一方、アメリカはトランプ大統領のメッセージに従ってファイティングポーズをとり続けるということならば、日本の処し方も変わらざるを得ないでしょう。右にアメリカ、左に中国がいて、両者が激しく対立しているならば、日本はその他の国々を巻き込んでグローバルな協力を呼び掛け、そこに米中を逆に引き込んでいくといった構図ができれば、これもまた日本にとってのチャンスになるかもしれません。

EU と協力してアメリカからの圧をかわすという考え方もあるのですが、こと AI に関しては規制色の強い「EU AI 法」が邪魔をして動きが取りにくい。パートナーとして誰と組むかによって、理想的な AI ガバナンスを目指すプロセス、経路が変わらざるを得ないのです。しばらくは過渡期の混乱が続くだろうという感じがしています。

## 日本の勝ち筋はどこに？

**菊池** AI ガバナンスの枠組みの構築に日本が貢献していくための課題は何でしょうか。

**飯田** 国内の AI の利活用です。これが進んでくれないと国際的にリーダーシップを取ろうにも限界があると感じています。「日本は AI ガバナンスで頑張っているけど、どんな AI を作っているのか？」とよく聞かれますが、残念ながら AI 開発で日本は後れを取っています。せめて利活用でユニークな事例を作り出し、世界と肩を並べることが、AI ガバナンスで日本が存在感を示せるかどうかの拠り所なのです。AI ガバナンスの取組の究極的目的は「安心安全に AI を開発、利用する環境を整え、AI による社会経済の発展向上を促進する」ことにあり、日本が目指すべきものであると私は考えていますので、そのためにもイノベーション促進的な枠組みが成果を上げることで日本が何らかのかたちで存在感を示すことは大切だと思っています。

残念ながら現時点では、現実的に依って立つ基軸は広島 AI プロセスぐらいしかなく、日本が AI ガバナンスで貢献できるかどうかはここにかかっています。それゆえ、同僚たちはとにかく広島 AI プロセスを発展させようと懸命に頑張っているのですが、私はもっと広い視野を持つべきだと注意喚起しています。いろいろな議論の場、それはガバナンスに限らず、普段から AI エコシステムに関わるあらゆる議論の場に参加し、幅広い観点を吸収して広島 AI プロセスの発展に反映させなければならない。そうしないと広島 AI プロセスだけにこだわり、環境や議論の流れの変化にも置いて行かれ、結局は広島 AI プロセス自体も無用のものになってしまうかもしれないからです。

ただし、日本の中立的な立ち位置や柔軟性のある対応力は、今のような国際情勢の中で本当に強みだと思います。フレンズグループに入りたいという中近東の国々の大使や担当者たちにこう言われたことがあります。

「素晴らしい仕組みなのでぜひ入りたい。しかし、あなたが G7、G7 と繰り返すたびにアメリカやイギリスが脳裏に浮かんできて警戒してしまう。日本がリーダーシップを取っているプロジェクトであり日本と一緒にやろうと言ってくれた方が、みんな安心して入れる」

日本の国際的リーダーシップの取り方を言い得ていると思いました。みんなの橋渡し役、仲介役になって仲間を作っていく。こうすべき、こうでなければならないではなく、柔軟性があって、包容力があって、みんなと一緒にやっという、ふわっと優しい枠組みがつかることができれば、そこに日本の勝ち筋が見えてくるのではないかと思います。

最近心配しているのは、各国が「ソブリン AI をつくる」ということを言い始めていることです。各国・地域の言語や文化に基づいた AI をつくるという考えは分かりますし、言葉そのものを否定する気はないのですが、「ソブリン」と言う言葉には、往々にして人も技術も内に抱え込んで、排他的、閉鎖的に進める意図が隠されていることが怖いのです。それは「ソブリン・データ」という発想につながっていき、クローズドな AI エコシステムの乱立につながる恐れがあり、とても危険です。

もう一つの心配は、中国による発展途上国の囲い込みです。今、中国の AI 戦略の中心は国際協力です。途上国に対して、人、技術、資金を含めた協力提供を積極的に喧伝しています。途上国側は、元々デジタル化で先進国に搾取されたという被害者意識が強いですし、AI でまた格差が広がるのではないかという警戒感もある。そもそも、AI どころかインターネットさえない場所もある。どうしても守りに入りがちなのですが、中国はそこに付け入り、笑顔で手を差し伸べて中国 AI の勢力圏の拡大を狙っているのです。しかし将来、知らないうちにあらゆるデータが中国政府の手に渡ってしまう。これは恐怖の構図だと思っています。

**菊池** AI ガバナンスの舞台裏での日本の貢献について具体的にお話いただき、まことにありがとうございました。

## 飯田 陽一 / Yoichi Iida

総務省 参与（前・情報通信国際戦略特別交渉官）

東京大学経済学部卒業後、1988年、旧郵政省入省。91年～93年、OECD事務局 ICCP（情報コンピュータ通信政策）課勤務。98年～01年、在ドイツ日本大使館一等書記官に就任。15年、総務省情報通信政策総合研究官、23年～25年（4月末まで）、総務省情報通信国際戦略特別交渉官。25年5月より総務省参与。2016年にG7情報通信大臣会合WG議長、19年にG20デジタル経済タスクフォース共同議長、23年にG7デジタル・技術大臣会合WG共同議長および「広島AIプロセス」WG議長を務めた。

インタビュー実施：2025年9月26日

# 作れなければ論じられない

喜連川 優

情報・システム研究機構 機構長

谷脇 AI法の成立を受け、今後、AI基本計画の策定など具体的な議論に入っていきます。日本はどの方向に進むべきなのか、どのような全体像を描くべきか、そのためのヒントをお伺いしたいと思います。

## 基礎研究なきガバナンス論は空回り

喜連川 ちょっと自慢してもいいですか（笑）。NIKKEI Digital Governance が米 Weights and Biases と共同作成している主要 AI モデルの性能比較データベース「AI Model Score 100<sup>1</sup>」で、僕たち、国立情報学研究所(NII)の「LLM-jp-3.1」が総合 21 位、日本勢でトップにランキングされたんです(NIKKEI Digital Governance の AI モデルランキング「AI Model Score」で「LLM-jp 3.1 8x13B」が日本製のモデルの中で第 1 位に<sup>2</sup>)。

2022 年 11 月に OpenAI が ChatGPT を正式ローンチし、生成 AI ブームが巻き起こりました。「なんや、このペラペラ喋る気色悪いものは!？」（笑）と戸惑っていたのですが、NII 所員に聞いて回っても「よく、わからない」という答えばかりです。当時は OpenAI が生成 AI の作り方に関する情報をかなり公開していたので、「分らんやったら作ってもうたらええやんか。モノマネだと言われてもかまわへんやないか」と発破をかけ、所内にプロジェクトが立ち上がったのです。僕自身は翌 2023 年 3 月末で NII の所長を退任することになっていたのですが、動かずにはいられなかった。

僕が言いたいのは、AI とは何かを理解していない人たちばかり、AI を作ることもできない人たちばかりで、「AI ガバナンス」を議論することなんてできますか、ということなんです。他人が作った AI プロダクトを評価したり、使い方をマスターするというのではなく、AI そのものを理解することが重要です。言い換えれば、基礎研究を抜きにしてガバナンスを論じても空回りしてしまうと思います。

今回の取り組みは、国費を使った「ソブリン AI」です。ソブリンということには苦い思い出があって、経済産業省が 2007 年に立ち上げた「情報大航海プロジェクト」という国産検索エンジン開発プロジェクトでは、100 億円以上が投じられましたが、メディアからは「国税の無駄づかい」「Google に勝てるわけがない」と厳しい批判を浴びました。しかし、今になって思い返しても、あれはやっておくべきだったし、やっておいて良かったと心底から思います。自分たちで学んで、考えて、作ってみて初めて分かることがたくさんあるからです。冒頭のランキングの件は、順位がどう優劣がどうということよりも、自分たちの試行錯誤や取り組みが的外れではないことを確認できたことに大きな意味があると思っています。

LLM（大規模言語モデル）という得体のしれない巨大ニューラルネットワークが、人間にとって制御可能なかどうか——。そういう基礎研究、根本的な研究をしっかりとやるべきというのが、僕の意見です。

1 <https://vdata.nikkei.com/prime/digital-governance/ai-model-score/>

2 <https://llmc.nii.ac.jp/topics/nikkei-digital-governance250627/>

## AI エージェントは RPA の二の舞に？

**谷脇** その問題意識、AI ガバナンスを論じていく上でとても重要な示唆を与えるものだと思います。別の視点から伺います。コンピュータの世界には「集中と分散」という議論があります。集中型の大型コンピュータから分散型の PC の普及、さらにクラウドのような集中型から分散型のエッジコンピューティングのような変遷です。AI の場合、LLM はこの「集中と分散」という文脈の中でどのようにとらえるべきでしょうか。

**喜連川** どちらの側面もあります。LLM は集中型ですが、最近の流行語である AI エージェントは分散型ととらえられます。かつてのコンピュータシステムがそうであったように、AI も非常に速いペースでその形態を変えているのです。

AI エージェントの分散化が行き過ぎると心配なのは、RPA（ロボティック・プロセス・オートメーション）の導入で起こったような混乱が AI でも起こるのではないかということです。RPA は定型的・反復的業務を自動化・省力化できるということで大流行し、大きな大学では RPA を 5000、1 万と作った。その結果、どうなったかというと**管理・メンテナンスが困難な複雑怪奇な寄せ集めシステム**になってしまいました。

AI エージェントでも同じような現象が起きるかもしれません。SLM（小規模言語モデル）がエージェント分野に特化して多数作られていったとき、最初は良くて、長期的に維持・運用し続けられるのか。AI エージェントを導入するのは、コンピュータシステムの専門家ではなく、AI 研究者や実務家ですから、維持管理のところまで考えが及ばない可能性もあります。

## 日本の実力を直視して、何を究めていくか

**谷脇** 先ほどの基礎研究をしっかりとということに関連するのですが、国は AI 分野にもっと投資すべきではないでしょうか。半導体に何兆円も投じていますが、それだけで大丈夫でしょうか。俯瞰的なアプローチが必要な感じもしています。

**喜連川** ご質問の含意は、日本が勝てたのはメモリーだけなのに苦手なロジック系に挑んで勝ち目はあるのか、では AI への投資はどのように行えば良いのか、と言い換えることができます。前者についてはぜひ成功してほしいと願っていますし、後者については自らを省みて戦略的に行う必要があると思います。

いずれにしても、しっかり認識しておくべきことは、日本は論理的に複雑で新奇なものをインベント（発明、創造）することが、少なくともこれまでは必ずしも得意ではなかったということと、責任の所在を曖昧にしてはいけないということの 2 点です。

近代における日本の成功・発展というのは、米欧が発明した技術や方法を学んで、真似て、改善を繰り返すという辛抱強い努力の賜物だったと言って良いのではないのでしょうか。コンピュータ然り、自動車然り、家電然り、半導体然り、そして、生成 AI 然りです。最初の「発明」の段階から日本が道を拓いたという成功事例は考えても



あまり思い浮かばない。1980年代に「Japan as No. 1」という言葉が流行りましたが、たまたま時代の巡り合わせが良かった“まぐれ当たり”だったのではないか。それは、その後の“失われた30年”の凋落を見れば自明ではないか——。年齢を重ね、過去を冷静に振り返ってみると、日本の実像がそんなふうに見えてくるのです。悲観ばかりでは前に進めませんが、驕り高ぶれば道を誤ります。**日本の実力を踏まえた上で「何を究めるのか」**を定めることをしないと、いくら金を注ぎ込んでも同じ轍を踏むことになります。

もう一つは、**政策とその結果に誰が責任を持つのか**を曖昧にしないということです。NIIのLLMプロジェクトについては、200億円以上の公費が投入されています。その結果責任を負うのは言い出した僕自身です。大学や組織の壁を越えて皆さんと一緒にやりましょうと声がけして、日本のトップクラス研究者200人くらいに参画してもらいました。アメリカの企業が作ったものに依存するのではなく、自分たちで考えて、作って、工夫して、真理に迫るというアティテュード（姿勢・考え方）で向き合いました。

僕はやってみて本当に良かったと思っていますし、参画してくれた皆さんも同じ思いだという自信があります。そうした研究者の熱意に加えて、政策担当者の敏感なセンスと俊敏なレスポンスが絶妙に噛み合った。これからのAI国家戦略を定めていくための礎の一角くらいにはなったんじゃないでしょうか。

ただし、**もし失敗したら、何が悪かったのか、どうすれば良かったのかを徹底的に究明して、次に生かす。それが責任を取るということ**です。誰も責任を取らない構造を放置していたから“失われた30年”などということになってしまった。僕は、そこを強く懸念しています。基礎研究への国家予算はどんどん減らされ、応用研究にシフトしてきましたが、それで日本は強くなったのか、このままでAI時代を生き抜けるのか。今こそ考え直すべき時だと思います。

**谷脇** 政府の「基本計画」というものは、「研究開発予算を倍増」「全国で実証実験100件」など、KPIの分かりやすさを追求した“お決まりメニュー”に陥りやすいところがありますが、何を実現するために何をどのようにやるのかという点を明確にすることが重要ですね。

**喜連川** AI基本計画を書く人にも覚悟が必要でしょうね。“お決まりメニュー”ではあまり役に立ちそうにはないですから。

学問の領域では、**研究スタイルそのものがAIの登場によって様変わり**しています。サイエンス領域に特化したファウンデーションモデルが数多く登場していて、テーマを入力すると、関連論文が一覧できるばかりか、研究活動が厚い領域と薄い領域が見えてくる。以前は、大量の論文を一生懸命読み込んで「この領域の研究が抜けている！」みたいなことを自分で見つけるしかなかった。その膨大な作業がAIで

ほぼ自動化できてしまうのです。また、AIに書かせた論文のレベルも急速に上がっています。まだ学会などではデスク・リジェクト（却下）されていますが、優秀な研究者が書いたものと見まがうような素晴らしい論文も出てきています。

もう、そういう時代になっているのです。研究のためにはAIを使わざるを得ない。本当に使わざるを得ないのです。よく「AIは思考の壁打ちに使えます」という人がいますが、壁を打ち破るくらいの勢



いで使い倒してほしいですね（笑）。AI 基本計画も AI に書かせてみたらどうでしょうか（笑）。

## 恐れだって来るものは来る

**谷脇** シンプルな質問です。先生は AGI（Artificial General Intelligence、汎用人工知能）や ASI（Artificial Superintelligence 人工超知能）の到来に恐れのようなものを感じることはありますか。

**喜連川** 恐れだって来るものは来るんですよ。百科事典が電子辞書に置き換わってしまったように、AI がいろいろなものを置き換えていくのは必然です。今は、その変化をどう受け止めるかを考える段階にあって、怖いか怖くないかを論じるのはあまり意味がないと思います。

**谷脇** AI ガバナンスの観点からすると、凄まじい勢いで進化していく AI を人間のスピードでガバナンスできるものでしょうか。

**喜連川** それは AI の問題というより、人間の問題ですね。悪意ある人間というのは、いつの時代も、どこにでも必ずいるもので、日本で規制をかければ規制の緩い第三国に移って悪さをするわけでしょう。日本は規制が厳しすぎるせいで、いろいろな分野で後れを取ったという過去を直視すべきです。その点、デジタル政策フォーラムが提言した通り、AI 法が規制法ではなく推進法となったことは喜ばしいことです。

そして、AI ガバナンスは必要ですが、それが日本の成長・発展に直接寄与するというものではないでしょう。AI 時代に日本がいかにか生き抜いていくのかについては、英知を結集して、ホンモノの戦略を練る必要があると思います。

**谷脇** 国のあり方と表裏一体の関係にある安全保障戦略の議論と似ていますね。貴重な示唆をいただき、ありがとうございました。

### 喜連川 優 / Masaru Kitsuregawa

大学共同利用機関法人 情報・システム研究機構 機構長

東京大学 特別教授

1983 年東京大学大学院工学系研究科情報工学専攻博士課程修了。データ工学の研究に従事。東京大学地球観測データ統融合連携研究機構長（2010-2013 年）、国立情報学研究所所長（2013-2022 年）などを経て、2023 年より情報・システム研究機構機構長。また 2021 年より東京大学特別教授、2023 年より東大デジタルオペザバトリ研究推進機構機構長。情報処理学会会長（2013-2014 年）、日本学術会議情報学委員会委員長（2014-2016 年）などを務める。

インタビュー実施：2025 年 7 月 30 日

# ジャーナリズムは slop に駆逐されるのか？

平和博

桜美林大学 教授

## 生成 AI によるジャーナリズムの民主化は歓迎

**菊池** 生成 AI の急速な進化・普及によって、ジャーナリズムはどのような影響を受けていますか。

**平** 生成 AI による「効率化」「低コスト化」「民主化」という大波が、ジャーナリズムを直撃しています。情報の収集から確認、編集、発信まで、これまで人が時間と労力をかけてやってきた作業が、定型的なニュースであれば、今はもうワン・プロンプトで回ってしまいます。

通常のニュースだけでなく、データジャーナリズムのようなテクノロジーを駆使した調査報道についても同様です。従来は情報分析の専門知識を持った記者、プログラマー、デザイナーらがチームを組んで行う、ややハードルが高いものでした。でも今は、情報公開で手に入れた資料を AI に読み込ませ、「この中で不自然な点を教えて」「注目すべきポイントは？」と聞くだけで、一定の示唆が得られる。専門知識がなくても直感的に取り組める環境が整いつつあります。その結果、ジャーナリズムの敷居が下がり、誰でもニュースの発信者になれる。偽情報・誤情報の氾濫が加速されるといった課題も抱えてはいますが、ジャーナリズムの「民主化」が進むこと自体は歓迎すべきだと思います。

一方、情報の価値とコストの関係がより鮮明になってきました。「情報は無料になりたがる」というのはスティーブ・ジョブズも愛読していた『ホールアース・カタログ』のスチュアート・ブランド編集長の言葉ですが、オンラインでたやすく取得できる情報、AI で簡単に加工できる情報が増え、コモディティ化が加速しています。ただブランドをあわせて、「価値ある情報は高価になりたがる」とも述べています。相対的に、現場に行かなければ取れない情報、オフラインにしか存在しない情報の価値が高まっています。いわば「足で稼ぐジャーナリズム」の価値に対する再評価です。また、情報の質を担保し、ユーザーにとって意味のあるカタチで届ける「編集機能」の重要性も高まっていると思います。

## データジャーナリズムに威力を発揮

**菊池** ジャーナリズムの実務に、AI はどのように浸透しているのでしょうか。

**平** OCR（光学文字認識）による資料のテキスト化・要約、取材音声や会議録の文字起こし・要約、大量のデータの分類・分析、スポーツ・企業決算・地震速報など定型的な記事の自動生成、さらには原稿の自動校正、SEO（検索エンジン最適化）タグや SNS 投稿文の生成、メールマガジンの作成などに幅広く利用され、主に労働集約的なバックエンド作業が大幅に効率化されています。

また、先ほどのデータジャーナリズムへの生成 AI の応用も確実に進んでいます。データジャーナリズムとは、大量のデータの分析から、見えなかったファクトやストーリーを見つけ出す手法であり、その源流を含めると半世紀以上前から試行錯誤が重ねられてきました。コンピューターの計算資源、専門知識を持つ

たエンジニアの助けも必要だったのですが、生成 AI を活用することで資源やさほどの知識がなくてもアプローチできるようになりました。データを投入し、上手にプロンプトを書けば、ある程度の成果を短時間で得られます。

調査報道の金字塔の一つがベトナム戦争をめぐる米国防総省の機密文書（ペンタゴン文書）に関わる報道ですが、告発者のダニエル・エルズバーグは 7000 ページあまりの文書を一枚一枚コピーして 1971 年にニューヨーク・タイムズ紙に持ち込みました。約半世紀後、租税回避行為を記した内部文書（パナマ文書）に関する調査報道では、2.6 テラバイトもの容量のデータとして持ち込まれ、クラウドコンピューターの並列処理にかけてデータベース化した上で分析しています。

最近では、AI を活用した報道がピューリッツァー賞（ジャーナリズム部門）を受賞しています。2024 年にはイスラエルによるガザへの大型爆弾攻撃を衛星写真のパターン認識によって特定したもの、2025 年にはイーロン・マスクの 150 万語におよぶ X 投稿からその思考と行動の変遷を分析したものなどがあります。今後、必ずしも高度な専門知識を持たないジャーナリストやグループが生成 AI を使って新たな挑戦を重ねていくでしょう。

ただし、「生成 AI によってすべての記事を作成する」という段階にまでは至っていません。生成 AI によって作成されるニュースコンテンツの「正確性」が担保できないためです。欧州放送連合（EBU）と英 BBC が主導した調査（News Integrity in AI Assistants, 2025 年 10 月 22 日）<sup>1</sup>では、主要な生成 AI（ChatGPT、Copilot、Gemini、Perplexity）に対し時事的話題について質問し回答を分析したところ、内容が誤っている、出典が明記されていないなど 45% に非常に重大な問題があったというのです。正確性はジャーナリズムの根幹であり、そこに対する懸念が解消されない限り、全面的な適用は難しいでしょう。

## 「公共圏」の土台がゼロクリックで崩壊？

**菊池** 生成 AI の登場でジャーナリズムの手法が少しずつ変わりつつある一方で、変わることはないジャーナリズムの根幹とは何でしょうか。

**平** 価値ある情報、信頼できる多様な情報を提供し続けることです。言い換えれば、立場が異なる様々な人たちが、意見を自由に交わすことができる場、すなわち「公共圏」の土台を支えることです。メディアには報道の倫理基準、トレーニングに裏打ちされた取材・編集のスキルや情報取り扱いのノウハウがあります。そんなメディアだからこそ担える役割であり、これらは生成 AI 時代でも揺らぐことのないジャーナリズムの基盤だと思います。

問題は、経済的に持続可能かどうかということです。多くの人々がネットやソーシャルメディアに溢れる無料のニュースで満足しています。サブスクの料金を支払ってまで報道機関のペイウォールの奥にある元記事を読みに行くという行動へ

<sup>1</sup> <https://www.ebu.ch/news/2025/10/ai-s-systemic-distortion-of-news-is-consistent-across-languages-and-territories-international-study-by-public-service-broadcast>



のハードルは、依然として高いと言わざるを得ません。

さらに、AI に聞けば欲しい情報をピンポイントで整理・要約して返してくれるので、それ以上いろいろなサイトをクリックして回らずに完結してしまう、いわゆる「ゼロクリック」という現象が広がっています。報道機関のビジネス環境は厳しさを増しています。

**菊池** メディアとプラットフォームはフレネミー(友を装う敵)の関係にありました。AI で状況は変わるのでしょうか。

**平** そもそも、メディアとプラットフォームの対立ポイントは、メディア側の主張としては、プラットフォームがニュースコンテンツに対して適切な対価を支払っていないという「ただ乗り論」でした。一方のプラットフォーム側はメディアのウェブサイト「送客」していると主張。相互依存関係にある両者ですが、その言い分は平行線のまま交わることはありませんでした。

この「ただ乗り論」はAI 時代にもそのまま持ち越されています。「ゼロクリック」が進めば、メディアのウェブサイトへの「送客」さえ無くなって単に「ただ乗り」されるだけになってしまうので、いよいよ看過できなくなっているのです。

ニューヨーク・タイムズやシカゴ・トリビューンが OpenAI や Perplexity を著作権侵害で提訴したり、日本の読売新聞社、朝日新聞社、日本経済新聞社が Perplexity を提訴したりといった最近の動きは、そうした危機感に根差しています。公正取引委員会が、生成 AI を使った検索サービスにニュース記事を無断使用することは独占禁止法違反の可能性もあるとして実態調査を始める方針との報道もあります。

その一方で、大手メディアとプラットフォーム事業者、生成 AI 開発企業との業務提携も進みつつあります。生成 AI 時代のフレネミー関係はまだら模様で、非常に流動的な様相を見せています。

## 著作権保護の強化、対価支払いの義務化も

**菊池** 健全なジャーナリズムを維持・確保するために、各国政府や国際社会が何らかの介入をすることが必要になるのでしょうか。

**平** 民主主義社会の機能としての公共圏を維持するためには、信頼と価値のある情報を担保する健全なジャーナリズム、健全なメディア空間が必要です。一定程度、政府の役割もあり得ると考えています。

EU (欧州連合) は 2019 年に発効したデジタル著作権指令 (デジタル単一市場における著作権に関する指令、略称: DSM 著作権指令) で、メディアに対してニュースコンテンツ使用対価の支払いに関するプラットフォーム事業者との交渉権を認めています。さらにオーストラリアやカナダでは、グーグル、メタを対象に、ニュースコンテンツ使用対価の事実上の支払い義務を課しています。

日本でも動きがあります。日本新聞協会、日本書籍出版協会、日本雑誌協会、日本写真著作権協会が 2023 年 8 月、「生成 AI に関する共同声明」<sup>2</sup> を発表しました。生成 AI に学習させるデータの著作権保護についての検討、具体的には著作権法の 30 条の 4 の改正を含めた著作物保護策を求める内容となっています。新聞協会はその後も、著作物保護強化を求める意見や声明を出しています。

公正取引委員会は 2023 年 9 月に、Yahoo! ニュースについて「ニュースメディア事業者との関係で優越的地位にある可能性がある」と指摘しています。さらにすでに紹介したように、2025 年 12 月には、生

2 [https://www.pressnet.or.jp/statement/copyright/230817\\_15114.html](https://www.pressnet.or.jp/statement/copyright/230817_15114.html)

成 AI を使った検索サービスでの報道機関の記事の無許可使用に関する調査に乗り出すことが報じられました。メディア側の自助努力やプラットフォーム側の自主対応では乗り越えられない課題に、さらに踏み込んだ整理が行われるのかもしれませんが。

ちなみに、新聞社やメディア企業を支えるための公的な財政支援というアイデアもあります。ただ、権力からの独立性や非党派性を損なわないものであることが必要です。さらに、メディアに対するトラスト（信頼）が低下傾向にある中で、世論の合意が得られるのか、という問題もあります。

## 「スロップ」まみれの社会にしないための議論を

多くの人々が、ニュースは Yahoo! が無料で提供しているもの、くらいにしか考えていません。ニュースはお金を払って読むものという意識は希薄化しています。メディアがどのようにしてニュースを作り出しているのかはおろか、信頼できるニュース、それを土台にした公共圏というものが社会にとってどれほど大切かということも理解されなくなっています。並大抵のことではありませんが、この共通基盤を再構築していくことが最も大切なのではないかと思っています。

このままでは、情報空間のどこを見回しても slop（スロップ、主に AI によって大量生成された低品質なデジタルコンテンツ）で溢れかえってしまうかもしれません。メディアも効率化と称してそれに乗っかってしまう一方、ユーザーは「ニュースは AI 生成で十分」と言い放つ社会、そして生成 AI で加速する偽情報・誤情報が蔓延する社会。それは極めて脆弱な社会だし、そのかなりの部分はすでに目の前にある現実です。そうではない未来を目指すには、AI 時代のレジリエント（強靱）な民主主義社会をどのように設計すべきか、その土台となる社会機能としてのジャーナリズムをどう持続させ、ニュースや情報の信頼をどう担保していくべきかという根本的な議論が不可欠だと思います。

**菊池** 流れを変えなければいけないと痛感しました。ありがとうございます。

### 平和博 / Kazuhiro Taira

ジャーナリスト、桜美林大学 教授（リベラルアーツ学群メディア・ジャーナリズムプログラム）

早稲田大学卒業後、1986 年、朝日新聞社入社。横浜支局、北海道報道部、社会部、シリコンバレー（サンノゼ）駐在、科学グループデスク、編集委員、IT 専門記者（デジタルウォッチャー）などを担当。2019 年 4 月から現職。2020 年 4 月～2021 年 3 月、国立国会図書館客員調査員。2022 年 9 月から日本ファクトチェックセンター（JFC）運営委員。2023 年 5 月から科学技術振興機構 社会技術研究開発センター（JST-RISTEX）プログラムアドバイザー。

インタビュー実施：2025 年 9 月 26 日

# AI時代の報道・・・語れないメディアの窮状

西田 亮介

日本大学 教授

## 「エモい記事」で悦に入っている場合なのか

**菊池** 2025年、「エモい記事論争<sup>1</sup>」に火をつけられました。『エモさと報道』(ゲンロン刊)<sup>2</sup>でも、新聞社の古い体質を憂いておられます。AI時代における「報道」はどうあるべきでしょうか。

**西田** 今のマスメディアは、AI時代の報道について語れるような状態にないと思います。

2020年代、特にここ1～2年で新聞、テレビというマスメディアのマスメディア性がよく分からなくなってきました。端的に言えば、これまでマスメディアとして担ってきた役割を果たせなくなりつつある。

かつての日本は新聞大国であり、新聞こそがマスメディアの代表でしたが、その土台は崩れ始めています。新聞の1世帯当たり購読部数は2008年に1を割ってから急減し、今では0.5を割っています(新聞の発行部数と世帯数の推移、新聞協会<sup>3</sup>)。2024年には、「全国紙」と呼ばれてきた毎日新聞と産経新聞が富山県での販売・宅配から撤退しました。夕刊については、各紙が全国で廃止に踏み切っています。記者の頭数はこの10年で約25%減り、地方支局の数や海外支局員も削減の一途です。広告費はインターネットへのシフトが加速し、2021年にはついにインターネット広告がマスコミ4媒体(新聞、雑誌、ラジオ、テレビ)の合計を超えました。

新聞社のリソースは細る一方ですが、偽情報への対処、ファクトチェック、働き方改革などやらなければならないことは山積みです。紙面を予定稿で埋め、ネットで使い古した記事を紙面で使い回すようなことが普通に行われるようになった。新聞紙面を何だと思っているのか、と言いたい。それに加えて、書き手一人称で、エピソード、ナラティブ主体の感想文みたいな「エモい記事」の横行が目につくようになったのです。何かを全力で批判するでもない、要はお涙ちょうだいの日常描写記事です。新聞はそんなものを書いている場合なのかと問題提起したら、自省するどころか猛反撃してきました(笑)。

新聞の「中の人たち」の意識はマスメディアだった頃のままで少しも変わっていないのです。AIの大波にメディア全体が飲み込まれないようにする方法を考えなければならないのに、報道の仕方も、記事の作り方も、考え方も、従来踏襲のまま。すべてが後手に回っていることに気づかず、自分たちこそが世の中の先を見据えていると勘違いしている。はっきり言って終わっています(笑)。もう少し穏やかに言っても、特に新聞社は日経と読売を除いて、ポイント・オブ・ノーリターンを超えている。

2025年にはマスメディア各社がようやくファクトチェックに本腰を入れ始めたことに業界の関心が集まりましたが、2010年代にやっておくべきでした。マスメディアのマスメディア性が自明で、メディア全体がまだ信頼されている時代ならファクトチェックがまともに機能する可能性は大いにあったのですが、すでに現在

1 <https://gendai.media/articles/-/136517>

2 <https://amzn.asia/d/06k9NnxQ>

3 <https://pressnet.or.jp/data/circulation/circulation01.php>

はオールドメディア全体が信頼されず、マスゴミなどと呼ばれてしまう始末です。状況が全く変わっているのです。報道機関には、信頼だけではなく、信頼感も重要であるということが理解されていません。

例えば「オリジネーター・プロフィール (OP) <sup>4</sup>」(インターネット上の情報の作成者・発信者をユーザーが確認するためのデジタル技術) によって「この記事は〇〇新聞のオリジナル情報です」というように出元を確認できたとしても、そもそもオールドメディア自体が信頼されていないのですから問題解決にはつながらないのではないのでしょうか。規制当局や業界あげて OP 一辺倒ですが、2020 年代においてはもはや実効性がないと考えています。

ファクトチェックにしても、信頼されていないオールドメディアがそれをやったところで人々の納得や安心を得られるはずがありません。メディアが重要な情報をつかんで発信しても、大衆は「オールドメディアの情報は鵜呑みにできない」と受け止めてしまうので、その情報がトリガーとなってむしろ社会の分断や対立を煽ってしまうかもしれない。非常に深刻な状態にあると思います。

旧来型マスメディアの代わりに信じられているのは、堀江貴文さん、中田敦彦さん、ひろゆきさんといった、YouTube などで観て慣れ親しんでいるインフルエンサーたちです。100%信じているわけじゃないにしても、人々はネットで慣れ親しんだ彼らが言うことならまずは聞いてみようと思うのです。もはや、新聞、テレビ、Yahoo! ニュースではないのです。新聞やかつてのマスメディアはその現状に真剣に向き合うべきであって、エモい記事など書いて悦に入っている場合じゃない。

テレビも 5 ~ 10 年遅れで新聞と同じ衰退の道を歩んでいます。テレビ局の最近の雰囲気は新聞の 15 年前、2010 年頃とよく似ていて、「個人ではテレビを見ない社員」が増えてきました。テレビのデジタル化は完全に出遅れていて TVer 一本足状態です。負けを認めて自省しようとしなのは新聞と全く同じ。特に地方ローカル局の弱体化は顕著で、自主製作番組比率は 1 割くらいしかありません。コンテンツこそが価値なのに、番組制作能力は弱る一方です。

しかも、民放は「自分たちにはリソースが少ないのだから報道は NHK におまかせ」と初めから降参している節がある。民放はキー局と地方局が系列を組んで報道ネットワークをつくっていますが、異なる会社の集合体ですから脆弱さがある。ひとつの組織で日本全国をカバーしている NHK とは状況がかなり違うことは確かなのですが、「報道は NHK 頼み」という空気さえ感じることは違和感を覚えます。でも、放送法改正では明確に NHK 弱体化の旗を振りました。NHK のウェブニュースは減りました。民業圧迫を声高に主張してきましたが、その後、報道を充実させたり、業績が回復した社はあるのでしょうか。

## 半世紀以上変われなかった日本の「報道」

**菊池** マスメディアには自ら変わる気はないのでしょうか。

**西田** 2025 年は、新聞・テレビ各社が「政治報道」の体制を変えると横並びで宣言した年でした。兵庫知事選、東京都議選、参議院選挙の結果を予見できなかった

4 <https://originator-profile.org/ja-JP/>



たこと、「選挙の公正」を過度に意識するあまりインターネット空間に不正確な情報が蔓延することを許したという反省からです。日本新聞協会は1966年の新聞協会声明（公職選挙法第148条に関する日本新聞協会編集委員会の統一見解<sup>5)</sup>）を引っ張り出し、選挙運動期間においても自由活発な報道ができるとする声明を発表（インターネットと選挙報道をめぐる声明<sup>6)</sup>）。テレビ各社も選挙運動期間中においても批判すべきは批判するという態度を表明しています。

しかし、結局何か変わったのか、何も変わっていないんじゃないか、少なくとも読者や視聴者の多くは変化を認識・体感できていないという認識です。「変える」と宣言したのに変われないというのは、とてもしんどい状況だと思います。1966年の声明を持ち出してくるのは良いとして、新聞はこれまでの約60年間、一体全体何をやってきたのか。66年声明を阻むボトルネックは何だったのかという反省も何も無いまま、単に言うてみただけという印象しかありません。

テレビでは、特に総務省の目が届きにくいからではないかと疑っているのですが、関西の放送局をはじめ各局で、報道番組なのか政治バラエティショーなのか一般視聴者には区別しがたい「情報番組」が台頭し、真偽の不確かな情報を含めて取り上げて面白おかしくセンセーショナルに大衆を扇動するというのが近年の流れになっています。その一方で、「インターネットの情報は信頼できない、マスメディアの信頼を取り戻す」などと主張しても説得力を持つはずがないでしょう。番組編成全体で信頼と信頼感を取り戻す打ち手が必要なはずですが、やってることと言っていることが違うんです。そのことに対する反省もないのがテレビの今です。

NHKは頑張っていると思いますが、放送法の制限と業界の足の引っ張り合いの中で力を削がれ、日本の報道は傷む一方です。世界に目を向けてみると、例えばイギリスでは、2030年代には地上波放送を廃止してインターネット放送に全面移行するということが一つの選択肢として真剣に議論されている（Future of TV Distribution、英国情報通信庁、2024年5月<sup>7)</sup>）。ドイツやフランスも同様ですが、日本ではそうした議論が全くありません。

その背景の一つに、インターネットを広く使えるようになってから30年以上、オピニオンメディアは出てきましたが「報道」を担うスタートアップが日本では全くと言っていいほど出てこなかったという事実があります。それは我々の社会の貧しさの要因でもあり、傷んでいくばかりの旧来マスメディアに依存せざるを得ない状況を生み出しています。

## 政策的介入の劇薬でメディア再生を

**菊池** メディア任せでは限界があるように感じます。

**西田** 業界には、メディアの保護や振興を政府が推進することに対する拒絶反応が強くあります。もちろん理解しますが、このまま放置していたら旧来メディアは衰退し、絶滅してしまうかもしれない。元気のよい新興の報道メディアも一向に出てこない。AIが「トラストな情報基盤」を代替することも当面は難しい。ご指摘の通り、事業者任せだけではトラストな情報基盤、健全な情報空間を維持できなくなる恐れが

5 [https://www.pressnet.or.jp/statement/report/661208\\_99.html](https://www.pressnet.or.jp/statement/report/661208_99.html)

6 [https://www.pressnet.or.jp/statement/report/250612\\_15908.html](https://www.pressnet.or.jp/statement/report/250612_15908.html)

7 <https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/269636-call-for-evidence-future-of-tv-distribution/future-of-tv-distribution-report-to-government.pdf?v=344045>

大きくなっています。

それは社会全体にとって極めて大きなマイナスです。いよいよ、補助金や振興策など政府の介入が必要な局面が来ていると考えています。それほど事態は深刻なのです。

ただし、政策的支援を受けるためにはオールドメディアが「トラストな情報基盤」としての信頼回復に努めなければなりません。そのうえで、それこそ民業としての限界なので支援が必要だという構図が重要です。ところが、一番厄介なのは、オールドメディアは自分たちが人々から信頼されているといまだに信じ込んでいることです。そのズレた意識を修正しなければ、AI時代の報道の在り方という重大テーマについて語ることもできないという、この窮状から抜け出せないでしょう。

**菊池** 政策的介入という劇薬を投入しなければならないところまで来ているという警鐘ですね。まことに、ありがとうございました。

## 西田 亮介 / Ryosuke Nishida

日本大学 危機管理学部危機管理学科 教授

東京工業大学 リベラルアーツ研究教育院 特任教授

1983年京都生まれ。慶應義塾大学総合政策学部卒業。同大学院政策・メディア研究科修士課程修了。同大学院政策・メディア研究科後期博士課程単位取得退学。同大学院政策・メディア研究科助教(有期・研究奨励Ⅱ)、(独)中小機構経営支援情報センターリサーチャー、東洋大学、学習院大学、デジタルハリウッド大学大学院非常勤講師、立命館大学院特別招聘准教授を経て、2015年9月から東京工業大学大学マネジメントセンター准教授。2016年4月より東京工業大学リベラルアーツ研究教育院准教授。2024年から現職。東京大学公共政策大学院非常勤講師等を兼任。博士(政策・メディア)。専門は公共政策の社会学。

インタビュー実施：2025年11月19日

# AI が揺さぶる安全保障の基盤

大澤 淳

中曽根康弘世界平和研究所 上席研究員

## 社会の集合知としての AI を守れ

谷脇 AI が社会、文化、産業、経済に与えるインパクトについては多くが語られていますが、いよいよ安全保障の中核にも入り込みつつあるという問題意識を持っています。安全保障の専門家としてどのように見えていますか。

大澤 AI の発達による安全保障面への影響は、情報戦・認知戦、サイバー攻撃、兵器運用の三つの領域で既に顕在化しています。

個別の話に入る前に、AI によって人々の認識がどのように変わっているのかという大前提を押さえておく必要があります。それは、人々が知識を獲得する方法が大きく変わってきたということです。

インターネット以前、知識の源泉は百科事典や専門書であり、図書館が社会の集合知が結集する場でした。何か知りたいことがあれば、図書館に行って本を選び、紙に書かれた知識を引き出すというやり方です。

インターネットの時代に入ると、検索によって知識を引き出すようになり、Wikipedia のようなデジタルの百科事典も急速に整備されていきました。社会の集合知は図書館からデジタル空間への移行が進んだのです。

そして今、AI の登場によって次の段階に入ろうとしています。人々は検索ではなく、AI に聞き、AI から直接答えを得るようになっていきます。検索では関連する情報の一覧を得られますが、AI は答えそのものを返してきます。AI が返す答えがそのまま「正しい知識」だと受け入れられることも少なくないでしょう、その傾向はますます高まっていくでしょう。つまり、AI が内部に持っているデータや学習結果が、社会の集合知として機能し始めているのです。

そうなってくると、ちょっと困ったことが起きてきます。AI の学習データを意図的に間違った学習（情報汚染、コンタミネーション）をさせれば、社会の集合知そのものを書き換えることが可能になってしまうのです。これは従来の情報操作とは質的に異なるものです。

例えば日中間の歴史認識については、学術的な議論を積み重ね一定の理解が形成されてきました。異なる主張があっても、論文や史料、専門家の議論を通じて検証され、少なくとも何が争点かということは共有されてきました。中国は日中戦争時に南京で 30 万人が虐殺されたと主張しますが、日本は当時の南京の人口は数万人程度なのでそのような大虐殺はあり得ないと反論しているというのは一例です。検索すれば、両方の主張が返ってきます。

ところが、AI 時代は、AI が学習した内容が正しい答えとして返ってきます。中国政府の主張をニュースサイトやそれらしく見える情報源から大量に学習させれば、AI に「南京大虐殺とは何か?」と聞いたとき、「30 万人が虐殺された」という中国政府の主張があたかも事実であるかのように返してくる。既に、そういう工作が情報戦の一環として行われ始めています。単なる情報操作の域を超えた、集合知を標的とし

た認知戦です。

もっと卑近なマーケティング領域にも影響は及びます。飲食店の集客では食べログのようなレビューサイトで上位に掲載されるためにユーザー評価の星を増やすことが優先課題でしたが、これからはAIの回答の中に入れるか、AIの回答の上位に食い込めるかが勝負になります。人々の消費行動がAIによって左右されてしまうのです。

AIを騙すために誤ったあるいは偏ったデータを学習させるという悪意のある行為から我々が使うAIを守ることも、安全保障上、社会生活上の新たな課題として浮上しています。

## AI活用で激しさを増す情報戦・認知戦

**谷脇** 同様の懸念を「アルゴリズム化されたネットワーク<sup>1</sup>」というコラムに書きました。これまでは人と人がネットワークを介してつながる構造でしたが、人と人の間にAIが入ってきたり、AIとAIが直接つながって会話したりするようになっていきます。人と人のコミュニケーションの場合、議論を重ねることによって妥協や新たな気づきに至り、より多くの人々が合意できる中間解（コンセンサス）を形成することができました。しかし、妥協や曖昧さを許さないAIが入ってくると、主張がどんどん先鋭化して落としどころが見つけれないという状態に陥るのではないかと懸念しています。安全保障の観点からは抑止戦略が効かなくなって暴発の恐れさえあるのではないのでしょうか。

**大澤** 今のところ、AIの軍事利用は自動火器管制や意思決定支援など部分的にところにとどまり、戦略的判断は依然として人間が担っています。AIの暴走によって抑止効果が崩壊するようなことはないと思います。

ただし、戦場の最前線では無人兵器の増加によって状況が変わりつつあります。ラストワンマイルでの攻撃判断をAIが担う場面が増え、国際法上、倫理上の問題が現実の課題として浮上してきています。そして、先ほども述べたように、より深刻なのは「情報空間」での戦いがあからさまに激しさを増していることです。昨今の日中関係の悪化を見ても、中国国内の対日感情、日本国内の対中感情がAIによって増幅され誘導されようとしており、危うさを感じています。



**谷脇** 日本としても、認知戦、情報戦への備えを本気で考えるべき段階に来ていますね。

**大澤** おっしゃる通りです。AIがからむことで、これまでにあった情報操作の手法が大きく変わってきています。

この十年、ロシアを起源として、情報戦、認知戦、影響工作と呼ばれる手法が欧米の情報空間で顕在化してきました。ヨーロッパではFIMI（Foreign Information Manipulation and Interference）という概念で整理されています。

初期は、トロール部隊と呼ばれる単純な方法でした。英語ができるロシア人を低賃金で雇って24時間

1 <https://www.digitalpolicyforum.jp/column/251106/>

3交代でSNSに大量投稿を行ない、アメリカの情報空間をロシア寄りに誘導するというものです。しかしこれはアメリカの情報機関による特定・摘発が進んで廃れました。

今の主流は、相手国民の発言を影響工作に活用する手法です。アメリカ国内のインフルエンサーの中から、親ロシア的、あるいは自国政府に批判的な言説を発信する人物を選び、その投稿を生成AIのBot群で増幅するというやり方です。直接お金を渡して工作をしているわけではありません。インフルエンサーの発言がBotによって拡散され、トレンドに乗ることで、アテンションエコノミーの仕組みの中で結果的に収益が発生する、間接的に報酬が支払われる構造です。ロシア発の書き込みなら規約違反で止められますが、アメリカ人によるアメリカ国内からの発信なので、表現の自由の壁があって規制が難しい。

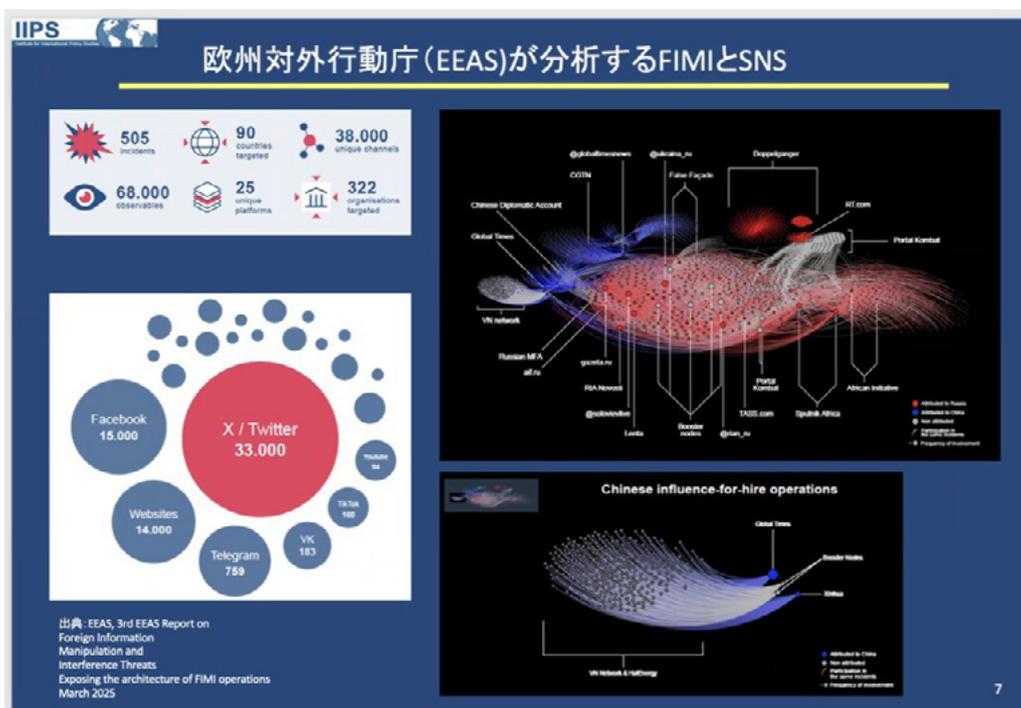
この新しい手法への対抗策としてアメリカ政府が狙いをつけたのは、インフルエンサー本人ではなく、増幅に使われているBotアカウント群です。これを司法省からXなどのプラットフォームに対して、自動生成・自動増幅の証拠と共に通知し、規約違反としてアカウント凍結を要請するのです。

例えば2025年7月に司法省は約1000アカウントを一斉凍結させたのですが、同一IPアドレス、同一ドメインから大量に作られたアカウントや異常な速度でリポストを繰り返すアカウントなどを、データ分析で洗い出しました。これは人手では不可能で、AI活用を前提としたビッグデータ分析の勝負になっています。そうした能力がなければ、情報戦の攻撃に対する防御が追い付かなくなっているのです。

## 影響工作の連鎖を断ち切る方策が必要

谷脇 同じことを日本でやろうとすると、かなり反発が出そうな気がします。

大澤 そうかもしれません。ただ、X、YouTube、TikTokなどは公開の掲示板であり、私信ではないので投稿情報を収集・分析すること自体は禁じられていません。実際、Xの規約では、人為的な増幅を目的としたアカウント、市民社会の意見形成を意図的に歪めようとするアカウントは凍結対象になっています。凍結されたアカウントが訴訟を起こすケースはほとんどありません。やっている側は悪意を自覚してい



ますから。

ちなみに、これは欧州対外行動局（European External Action Service: EEAS）が中国・ロシアによる情報戦を分析した年次レポート（3rd EEAS Report on Foreign Information Manipulation and Interference Threats、2025年3月）<sup>2</sup>からの抜粋です。2024年における90カ国にわたる505件の事案を分析すると、FIMIに悪用されているSNSプラットフォームとしてはXが圧倒的に多い。それにFacebook、ウェブサイト、テレグラム、TikTok、YouTubeが続きます。

右側上段にグラフ理論による分析図があって、赤がロシア系、青が中国系の影響工作で、中国系が増加中です。その図の左側の部分を拡大したものを下に貼り付けてありますが、公的な中国メディアであるGlobal Timesが発信した情報をBot群がブーストしているという分析がなされています。これは従来ロシアがやってきた手法ですが、中国もそれを真似てやり始めているようです。プラットフォーム側に規制を強化してもらう必要があると思っています。

**谷脇** 2025年の参議院議員選挙の際にも、Botによる偽情報、誤情報の拡散が指摘されました。日本でも認知戦対策は重要な研究課題ですね。

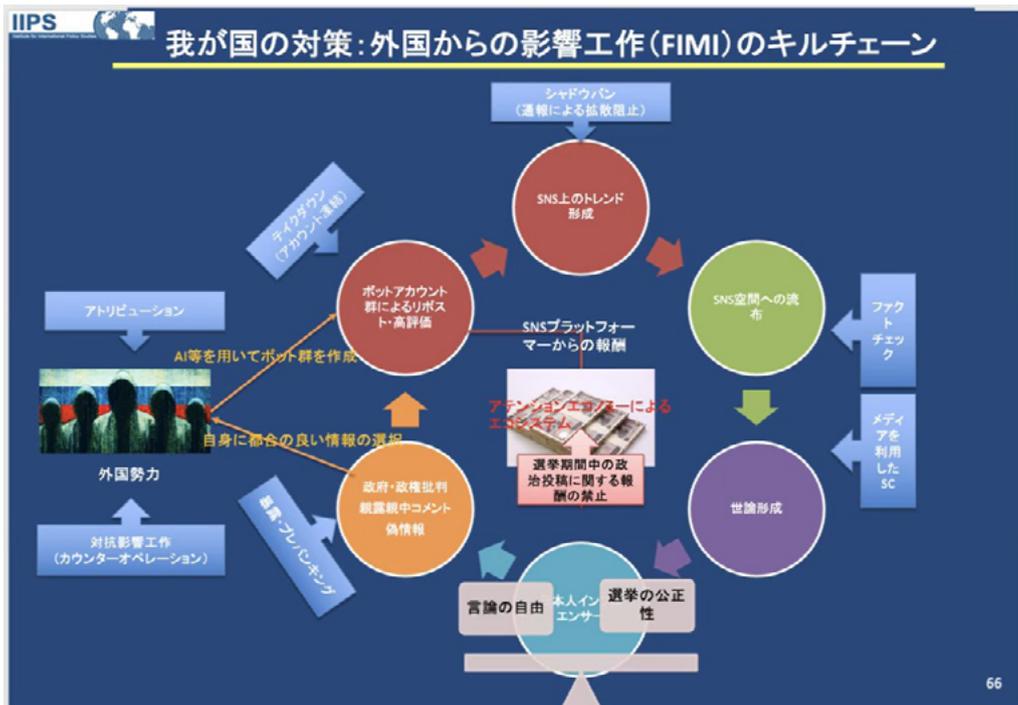
我が国の影響工作の対策の現状		
	日本	現状
メディアリテラシー教育	△	小学校学習指導要領で「情報活用能力（情報モラル）」は規定されているが、フェイクニュースを見抜く等のメディアリテラシー教育は規定されていない。
ファクトチェック	△ (3機関)	日本のファクトチェック機関は小規模であり、EUのような公的資金の投入がない。また、政府によるファクトに関する訂正も積極的には行われていない。
プラットフォーム規制	×	プラットフォーム規制関連で検討されている法案は現時点でない。
オフライン規制	△	政治資金規正法第22条の5で外国人・外国法人からの政治献金が禁止されている。
刑法的制裁	○	公職選挙法第142条の7に「公職の候補者に対して悪質な誹謗中傷をする等表現の自由を濫用して選挙の公正を害することがないよう、インターネット等の適正な利用に努めなければならない」との規定があり、罰則としては、虚偽事項公表罪、名誉毀損罪や侮辱罪が定められているが、選挙時のSNS投稿に関する起訴事例はない。
調査・モニタリング	△	一部省庁で所管内容と関連したモニタリングはあるが、一元的、体系的な収集分析体制は未確立。
戦略的コミュニケーション	△	偽・誤情報対応は中心には据えられていない。「原発処理水」等で外務省の運用実績があるのみ。

**大澤** はい。特に選挙は深刻です。

AIに限らないのですが、影響工作には多層的な対策が必要になっています。メディアリテラシー教育、ファクトチェックに加えて、プラットフォーム規制、オフライン規制、刑法的制裁を組み合わせで対策していく必要があると思います。

ただし、インフルエンサーの発言から政府批判や親口・親中のコメントを選択的に抽出して、それをAIを使ったBot群で増幅し、トレンドを作り、結果的にアテンションエコノミーでインフルエンサーに報酬が渡るという仕組みなので、その循環をどこかで断ち切らなければなりません。

<sup>2</sup> <https://www.eeas.europa.eu/sites/default/files/documents/2025/EEAS-3nd-ThreatReport-March-2025-05-Digital-HD.pdf>



現実的な対策として考えられるのは、選挙期間中に限って、特定候補や政治テーマに関する投稿への報酬支払いを制限するといった措置があります。1年365日常に規制するのは難しいとしても、選挙期間という「民主主義の基盤として最も重要な期間」に限定すれば、一定の合理性はあると思います。

**谷脇** 同感です。ただ、選挙期間前はどうか、という問題は残りますね。

**大澤** そこはまさに言論の自由と選挙の公正性、社会の安定性との比較衡量になります。だからこそ恒常的な規制ではなく、時限的・限定的なルール設計が現実解だと思います。あるいは、プラットフォーム側にインフルエンサーに支払う報酬について源泉徴収を義務付ければ抑止効果があるでしょう。法改正が必要ではありますが……。また、日本政府のデータベースが税務署、出入国管理局、情報当局とすべて切り離されているので突合して分析することができないという問題もあります。

**谷脇** それはできないですね（苦笑）。特にAIとの関連で他に注意すべき論点はありますか。

**大澤** 危惧しているのは偽ニュースサイト問題です。一見すると地方紙やニュースサイトのように見えますが、よく見ると違和感だらけの異様な日本語記事が大量に掲載されているのです。そこに5%ぐらいの割合で中国寄りの記事が紛れ込んでいます。

これを説明できる唯一の仮説は、AIに学習させるためのデータを集めて掲載しているということです。既存のクオリティペーパーの多く、アメリカならニューヨークタイムズやワシントンポスト、日本では日経、読売、朝日などは多くの記事を有料にしているためペイウォールを超えないと閲覧できないようにしているためAIが勝手に学習することができません。そこで、AIに学習させるためのだけのサイトをわざわざ作っていると考えられるのです。このような偽ニュースサイトについては、2024年に、トロント大学の研究機関がレポートを公表<sup>3</sup>したり、読売新聞が報道<sup>4</sup>したりしています。

例えば Fujiyama Times、Fukuoka-ken.com、共同ノート、NHK ニュースといったドメイン名や名

3 <https://citizenlab.ca/research/paperwall-chinese-websites-posing-as-local-news-outlets-with-pro-beijing-content/>

4 <https://www.yomiuri.co.jp/national/20241105-OYT1T50001/>

称を使い、あたかも正規のニュースサイトであるように見せていますが、全くの偽物です。ですが、AI がネット上のコンテンツを学習する際には真贋の区別がつかないので、日本語ニュースサイトに掲載されたデータとして読み込んでしまいます。同じような偽ニュースサイトを多数つくり、そこに偽ニュースを流せば AI の学習データを汚染できてしまうのです。その結果、AI に政治や歴史について尋ねると、歪められた答えが出典メディア付きで返されるという仕組みです。

トロント大学の CitizenLab の報告によれば、中国によって世界 30 カ国、123 以上のサイトをネットワークした偽情報ネットワークが構築されていて、日本は韓国に次いで主要なターゲットになっています。各国の認知空間を長期的にコントロールする高度な情報戦・認知戦であり、CitizenLab はこれを「Paperwall 作戦」と名付けています。

## ソブリン AI の議論、経済合理性よりも安全保障優先で

**谷脇** ここまでのお話を踏まえると、やはりソブリン AI やソブリンクラウドといった議論を避けて通れないと感じます。中国は 2023 年に国家データ局を設置し、データの活用やデータスペースの構築を含めて、非常に戦略的に動いています。日本はそこまで俯瞰的な戦略を描けていないように見えます。

**大澤** 中国は、AI をめぐる戦略をかなり明確に持っています。国内向けには、生成 AI の出力やデータに透かしを入れることを義務づけるなど、自国の認知領域を守る法制度を次々に整備している。一方で、国際的には「自由にやらせろ」という立場を取る。中国の認知戦の戦略文書を見ると、「自分たちの認知領域を守り、相手の認知領域を攻撃する」という発想が明確に書かれています。

そう考えると、日本語の偽ニュースサイトや学習データの汚染といった動きも、近視眼的、偶発的なものではなく、長期的な戦略的行動として理解できます。

**谷脇** 中国は国際標準の主導権争いでも非常に積極的です。かつて、ITU（国際電気通信連合）や ISO（国際標準化機構）などの会合に出席しているのは中国企業ばかりで、中国語で議論をしているような状況がありました。AI についても、国産の中華 AI を守りつつ、国外では自由という名を掲げて中国主権を握ろうとしている意図がありありです。その姿勢は一貫しています。

**大澤** そうですね。今、中国はデジタル透かしの国際標準を取ろうとしています。そうすれば中国で生成した AI 情報は真正情報としてコントロールでき、逆に欧米から中国に入ってくる情報は規制できるようにすることを見据えた戦略的行動だと思います。デジタル透かしの標準化を進めている W3C（ワールドワイドウェブコンソーシアム）の主導権はほぼ中国に握られている状況です。

中国の「中国製造 2025」という計画は有名ですが、これとは別に「中国標準 2035」という計画もあります。中国が世界の大国として君臨するために国際標準を重要ターゲットとして位置付けているのです。日本としては、国際ルール作りの場で知らないうちに中国に有利なルールができてしまうことがないように強く警戒しなければなりません。

そして、ソブリン AI、ソブリンクラウドは、日本の安全保障のために必要だと思います。日本人の集合知を適正に維持するという観点からも、国産 AI、国産データベースを確保し、それを日本人が使うようにしないと、日本の情報空間はどんどん歪められていってしまうでしょう。その兆候は既にあります。大きなリスクだと感じています。

**谷脇** デジタル主権、デジタル空間でのソブリンティについて十分意識した政策が必要ですね。

**大澤** 動画を生成するとか、創作活動に活用するというような一般的な利用はともかく、日本という国に行く末を考えたり判断したりするための知識を得たり、ニュースを読んだり、あるいは政治情報を発信したりするといった用途に使われる AI は、国産のものを使う必要があると思います。欧米製の AI の場合は、何らかのルールに基づいて学習データを適正にコントロールすべきです。

また、政府や国民のデータを預けるのは、日本国内で日本企業が構築・管理するソブリンクラウドで、セキュリティクリアランスを持った日本人が運営するということが最低限必要だと思います。経済合理性よりも国家安全保障を優先すべき問題です。

**谷脇** サイバーセキュリティの領域では、AI を活用してエクスプロイト（ソフト・ハードの脆弱性を狙った攻撃）が自動化されています。攻めも守りも AI エージェントの戦いということになれば、エスカレーションが止まらなくなってしまうのではと危惧しています。

**大澤** 100 パーセント共感します。システムにおける脆弱性の公表から、そこを目掛けてサイバー攻撃が始まるまでのタイムラグが短くなり、今や 24 時間を切っています。以前は大企業など金になりそうな標的を選別した「点の攻撃」が主流でしたが、最近は脆弱性があるところを片っ端から侵入してみる「面の攻撃」へと移行しています。守る側も「面の防御」が必要になり、対策コストがどんどん上がります。

このまま進めば、ネットワークを使うための総コストが上がり、DX（デジタルトランスフォーメーション）が遅滞することにもなりかねません。インターネットは危ない、ファクスで受発注した方が安全だし安上がりだというような先祖返りにつながるかもしれません。笑い話ではありません。これまでに築いてきたインターネットの信頼性が AI の登場で加速度的に崩れてしまう可能性があるのです。

## 「デジタルのベルリンの壁」に対するため横断的国家戦略を

**谷脇** AI と安全保障の文脈で、国際的な合意やルール形成によって歯止めをかけることは難しいのでしょうか。2024 年には欧州評議会が AI 関連で初となる国際条約（人工知能と人権、民主主義及び法の支配に関する欧州評議会枠組条約の署名）<sup>5</sup> を採択し、日本も署名しました。また、サイバー空間における国家間の攻撃・防御について国際法の適用や基準について明文化したサイバー戦争のルールブックとも言える「タリン・マニュアル」<sup>6</sup> もアップデートが重ねられています。しかし、いずれも万能ではないように感じます。

**大澤** なかなか難しいですね。中国・ロシア、グローバルサウス、西側先進国それぞれの考え方がバラバラなので、包括的な合意は期待しにくい状況です。結果として、通信や情報空間はブロック化され、インターネットのフラグメンテーションが進みつつあります。中国やロシアは既に国内だけでネットワークを維持できるシステムを整えつつあり、実験も行っています。

**谷脇** 「デジタルのベルリンの壁」という表現が現実味を帯びてきますね。

**大澤** はい、既に言論空間は事実上分断されています。中国はグレートファイアウォール（金盾）を介して国外から入ってくる情報をコントロールし、中国語の空間に西側の論調が入り込んでこないようにしています。ロシアも最近、欧米プラットフォームの SNS を国内で利用することを禁じたりしています。これは、ウクライナ戦争の影響が大きいようです。

5 [https://www.mofa.go.jp/mofaj/press/release/pressit\\_000001\\_01725.html](https://www.mofa.go.jp/mofaj/press/release/pressit_000001_01725.html)

6 <https://ccdcoe.org/research/tallinn-manual/>

こうした状況だからこそ、最低限、ソブリン AI やソブリンクラウドが必要なのです。メディア（報道）、広告・マーケティング、著作権の既存モデルは早晚成り立たなくなるでしょうから、価値と対価がきちんと循環するエコシステムの再設計も必要です。

**谷脇** 全体を俯瞰して国家戦略として考える必要がありますが、日本政府は縦割りが過ぎて横串を通すのが苦手です。

**大澤** そうですね（笑）。AI 時代の安全保障、AI 時代のサイバーセキュリティ、AI 時代の経済安保といった横断テーマは、単一の省庁だけで扱える代物ではありません。個人的には、縦割りの組織に横糸を通すのは、「人のネットワーク」しかないと思っています。泥臭いのですが、忘年会や新年会で一緒に飲むとか、同期入省・入社の人々のネットワークを大切にするとか……。そうした横軸のコミュニティをいかに形成するかは、日本の組織における古くて新しい課題です。意識的に紡いでいかなければいけません。

**谷脇** とても大切なことだと思います。大澤さんが所属されている笹川平和財団や、手前味噌ながら私に取り組んでいるデジタル政策フォーラムといったオープンな場でのコラボレーションが広がっていくと、分野横断的な議論に厚みを加えることができるかもしれません。

**大澤** そうしたコミュニティの中で政治家、官僚、学者、経済人、企業人の横のつながりを形成していきたいですね。それは AI にはできないことですから。

**谷脇** たいへん奥深いお話をいただき、ありがとうございました。

## 大澤 淳 / Jun Osawa

公益財団法人 中曽根康弘世界平和研究所 上席研究員

1971 年生。慶應義塾大学法学部 1994 年卒、同大学大学院法学研究科修士課程 1996 年修了（法学修士）。1995 年世界平和研究所研究員、2009 年同主任研究員、2014 年～ 2016 年内閣官房国家安全保障局参事官補佐（サイバー安全保障担当）、2017 年～ 2025 年中曽根康弘世界平和研究所主任研究員、2025 年より同研究所上席研究員。

インタビュー実施：2025 年 11 月 25 日

# AI時代の個人情報保護 ～新たな課題と不変の原則～

堀部 政男

一橋大学 名誉教授

## はじめに

個人情報保護とプライバシーの問題は、1970年代のコンピュータ導入期以来、技術革新と制度設計が相互に作用しながら発展してきた。初期の段階では、電子計算機による住民データの集中管理が市民社会に深刻な懸念を呼び起こし、プライバシー侵害や権力の濫用への不安が高まった。その後、OECD（経済協力開発機構）や欧州を中心に国際的な法制度・ルールが形成され、わが国も段階的に制度を整備してきた。



AI（人工知能）、IoT、ビッグデータ、さらに生成AIの登場と急速な発展は、従来のデータ保護法制が想定してこなかった新しい課題を突きつけている。アルゴリズムの不透明性や差別的バイアス、説明責任の欠如、そして大規模データ解析に伴う監視社会化への懸念等は、プライバシー法制にとって新たな挑戦を意味する。

本稿では、歴史的経緯を踏まえつつ、AI時代における個人情報保護の課題を簡潔に整理し、日本にとっての政策的示唆を提示する。

## 1. コンピュータ導入期と個人情報保護の萌芽

### 1.1 1970年代の社会的懸念

1970年代、電子計算機（コンピュータ）の導入は行政や企業の業務を効率化する一方で、住民の個人情報的大量に集中管理されることへの懸念等を招いた。特に日本では、自治労（全日本自治団体労働組合）をはじめとする労働組合が合理化・人減らしへの危機感を背景にプライバシー侵害の問題を強調し、議論を広める役割を果たした。

1975年、東京都国立市は全国で初めて「コンピュータ条例」と呼ばれる個人情報保護条例（国立市電子計算組織の運営に関する条例）を制定した。これは住民の個人情報が無制限に利用されることを防ぐものであり、その後、東京の特別区を中心に条例制定が広がり、地方自治体による草の根的な情報保護の制度化が進展した。

### 1.2 米国における「公正情報行動原則（FIPPs）の提案

米国では1960年代末からプライバシー権の議論が活発化し、1973年には、米保健教育福祉省（HEW）の諮問委員会が、**情報源の特定（Limits on the Collection）、情報収集の通知（Notice）、情報収集目的の明示（Purpose Specification）、情報へのアクセス（Access）、情報利用の制限（Use and Other Disclosure Limitations）**という5つの原則を提案した。これが、現在の「**公正情報行動原則**」（FIPPs：Fair Information Practice Principles）の基礎となった。これに基づき、米連邦政府は翌1974年、「**プライバシー法**」を制定した。

### 1.3 欧州における初期の法制化

同時期、欧州では州レベルや国レベルで個人データ保護法が整備され始めた。国家による監視の歴史的記憶もあり、個人の基本権としてのプライバシー保障を強調する傾向が顕著であった。ドイツ・ヘッセン州は1970年に世界初のデータ保護法を採択し、スウェーデン、フランスなどが続いた。これらの動きは、やがて欧州全体の共通ルールの形成へとつながっていく。

## 2. 国際的制度化の進展：OECD、EU、日本

### 2.1 OECD プライバシーガイドライン（1980年）

1980年、OECDは「**プライバシー保護と個人データの国際流通に関するガイドライン**」を採択した。FIPPsを国際的原則へと昇華させたもので、「**収集制限**」「**データ内容の正確性**」「**目的明確化**」「**利用制限**」「**安全保護**」「**公開性**」「**個人参加**」「**責任**」の8原則を示した。この勧告は法的拘束力を持たないが、個人情報保護法制の国際的な共通基盤となった。国際的なデータ流通と個人情報保護の両立を目指した点は特筆すべきである。

### 2.2 EU データ保護指令（1995年）とGDPR（2018年）

欧州連合（EU）は、1995年に「**EU データ保護指令**」を制定し、加盟国に個人データ保護法の制定を義務付けた。その後、2018年にはより強力な「**一般データ保護規則（GDPR）**」が施行された。GDPRは以下の点で画期的である。

- ・ 域外適用：EU域外の企業であっても、EU居住者のデータを処理する場合に適用される。
- ・ データ主体の権利：アクセス権、訂正権、削除権（いわゆる「忘れられる権利」）、データポータビリティ権などを明記。
- ・ 高額な制裁金：違反に対して最大で全世界売上高の4%または2000万ユーロの制裁金を科す。
- ・ GDPRは、プライバシーを「人権」として位置づけ、強力な規範的影響をグローバルに及ぼしている。

### 2.3 日本における制度化

日本では2003年に「**個人情報保護法**」（個人情報の保護に関する法律）が成立し、2005年に全面施行された。その後、複数回の改正を経て、2021年の改正では、個人の権利救済の強化、越境データ

移転に関する規律の整備、個人情報保護委員会の権限拡大などが行われた。これにより、日本も国際的な十分性認定を受け、EU との間で円滑なデータ移転が可能となった。

## 3. AI 時代における新課題と対応

### 3.1 バイアス・差別、説明責任、著作権侵害

AI 技術、とりわけ機械学習や生成 AI の進化は、従来のデータ保護法制の射程を超える新たな課題を提起している。第一に、**アルゴリズムによるバイアスや差別**の問題である。データに内在する偏りがそのまま機械学習モデルに反映され、不公平な結果を生み出す事例が報告されている。第二に、**説明責任の欠如**である。ディープラーニングを中心とする「ブラックボックス」モデルは、その予測結果や判断過程を人間が十分に理解・検証できない場合が多い。第三に、生成 AI に特有の課題として**学習データの正当性や著作権との関係**が浮上している。

### 3.2 ICDPPC の「AI 宣言」(2018 年)

国際的なデータ保護機関の会合である「データ保護プライバシー・コミッショナー国際会議」(ICDPPC: the International Conference of Data Protection and Privacy Commissioners) は、2018 年に「AI における倫理及びデータ保護についての宣言」(Declaration on Ethics and Data Protection in Artificial Intelligence) を採択し、AI が人権・プライバシーに与える影響を強調した。また遡る 2017 年には、「コネクテッドカー決議」(自動化・コネクテッド(接続)された車両のデータ保護に関する決議、Resolution on data protection in automated and connected vehicles) を採択し、AI/IoT 社会におけるデータ保護の重要性を確認した。

### 3.3 OECD の「AI 勧告」(2019 年)

OECD は 2019 年に「AI に関する理事会勧告」を採択し、5 つの原則を示した。

- ・ 包摂的な成長、持続可能な開発及び幸福
- ・ 人間中心の価値観及び公平性
- ・ 透明性及び説明可能性
- ・ 頑健性、セキュリティ及び安全性
- ・ アカウンタビリティ

### 3.4 G20 大阪サミットが「人間中心の AI 利用」を明記(2019 年)

2019 年の G20 大阪サミット首脳宣<sup>1</sup> は「人間中心の AI の利用」を明記し、OECD AI 原則の普及を支

1 [https://www.mofa.go.jp/mofaj/gaiko/g20/osaka19/jp/documents/final\\_g20\\_osaka\\_leaders\\_declaration.html](https://www.mofa.go.jp/mofaj/gaiko/g20/osaka19/jp/documents/final_g20_osaka_leaders_declaration.html)

持した。主要国間での AI ガバナンスの協調は、今後の国際的制度改革に大きな影響を与えるだろう。

### 3.5 EU の「AI 法」(2024 年)

EU は 2024 年に「AI 法 (AI Act)」を制定した。同法はリスク・ベースの規制を採用し、「容認できないリスク (Unacceptable risk)」「高リスク (High risk)」「限定的リスク (Limited risk)」「最小のリスク (Minimal risk)」の 4 段階に分類し、それぞれに異なる規制を課す。このアプローチは、GDPR の一般的データ保護規制を補完しつつ、AI 固有のリスクに対応する新しい規範的枠組みを提示するものである。

## 4. 不変の理念：プライバシー・バイ・デザイン

このように、AI の急速な進化と普及に対応するため、法制度やルールの整備・見直しが進められているところであるが、その一方、技術革新にかかわらず普遍的に有効性を維持する理念も存在する。例えば、1990 年代半ばの議論をあえて挙げるならば、その代表は「プライバシー・バイ・デザイン」である。同概念はカナダ・オンタリオ州の情報&プライバシー・コミッショナーで、私の友人であるアン・カブキアン (Ann Cavoukian) 氏が提唱したもので、システム設計の初期段階からプライバシー保護を組み込むことを要請する考え方である。

具体的には、「事前予防的であること」「デフォルトで保護されること」「設計段階から組み込まれること」「利益・利便性の両立」「ライフサイクル全体を通じての保護」「透明性の確保」「利用者中心」という 7 原則から成り立つ。

AI 時代において、この理念は新たな意味を持つ。例えば、機械学習の設計段階でデータ最小化を徹底すること、説明可能性を確保するアルゴリズムを採用すること、ユーザーが自らデータ利用をコントロールできる仕組みを設計に組み込むことなどは、まさにプライバシー・バイ・デザインの具体的実践と言える。こうした理念は規制の補完的基盤として、技術者・事業者・規制当局の間で共有される必要がある。

## 5. 日本への示唆

日本はこれまで国際的議論を踏まえ、個人情報保護法の改正や個人情報保護委員会の設置を進めてきた。しかし、AI の急速な展開により新たな課題に直面している。

特に、生成 AI の学習データに関する適法性確保、アルゴリズムによる差別の防止、説明責任を果たすための仕組みづくりは、国内外の信頼を確保するうえで不可欠である。

さらに、日本は経済社会のデジタル化を進めるうえで、「信頼に基づく AI (Trustworthy AI)」という価値観をいかに制度や実務に落とし込むかが問われている。これは単なる規制遵守にとどまらず、企業や行政が自主的に倫理的基準や透明性を確保し、市民社会との信頼関係を構築することを意味する。そのためには、産官学の協働によるガバナンス体制の確立や、グローバルな規範形成に積極的に関与する姿勢が重要となろう。

## まとめ

AI時代の個人情報保護は、歴史的経緯の中で蓄積されてきた規範的原則と、新しい技術に対応する柔軟な規制枠組みをいかに統合するかという課題に直面している。すなわち、技術変化への適応（adaptability）と、基本原則の堅持（normative stability）の両立こそが、今後の個人情報保護制度の方向性を規定する。

日本においては、国際的規範形成への積極的関与と、国内における信頼性の高いAI実装の両輪を通じて、AI時代にふさわしい個人情報保護の在り方を構築することが求められる。

### 参考文献

『情報通信政策研究』第3巻第1号 [特集] AI/IoT時代のプライバシー・個人情報保護 特別寄稿「AI/IoT時代のプライバシー・個人情報保護の新課題」、堀部政男（一橋大学名誉教授・元個人情報保護委員会委員長）

[https://www.soumu.go.jp/main\\_content/000656379.pdf](https://www.soumu.go.jp/main_content/000656379.pdf)

Ann Cavoukian, Privacy By Design 堀部政男/JIPDEC 訳編『プライバシー・バイ・デザイン』（日経BP社、2012年）

### 堀部 政男 / Masao Horibe

一橋大学 名誉教授

1962年 東京大学大学院修士課程（基礎法学）修了

東京大学助手を経て、一橋大学法学部講師、助教授、教授

1994年 一橋大学法学部長・大学院法学研究科長（1996年まで）

1997年 中央大学法学部・法学研究科教授就任

2004年 中央大学法科大学院（ロースクール）教授就任

2007年 中央大学定年退職

2014年 特定個人情報保護委員会委員長就任

2016年 個人情報保護委員会委員長（2018年法定任期5年終了で退官、初代委員長）

情報関係の著作多数。

インタビュー実施：2025年7月22日

# 構造変化モデルで AI 社会を見通す

國領 二郎

慶應義塾大学 名誉教授

## 局所ではなく全体構造をつかめ

菊池 AI が社会に対してどのような影響を及ぼしていくのかについて、國領先生はどのように見えていますか。

國領 経済学がマクロ、経営学がミクロとすれば、今回はその中間くらいの視点でお話したいと思います。

アプローチとしては「構造変化モデルで考える」ことが必要だと思います。社会と技術のシステムアーキテクチャがどのように変化するかという観点からの研究を、私はインターネットについてずっとやり続けてきました。それを AI に当てはめてみたい。

### 構造変化モデルで考える

従来は「生産性」「置き換え」に関心集中

- AIは人間を置き換えて労働を奪う？

より大きな変化は「構造変化」から生まれる

- 仮説：自動運転はサービスモデルで提供される？

生産性向上は全ての企業に当てはまるから競争優位には結びつかない？

今のところ、AI に対する関心は「生産性」と「置き換え」に集中しています。それはそれで大切なのですが、生産性の向上はすべての企業に当てはまるので競争優位には結びつかないのではないかと、既存の構造の中でサブシステムの効率を高めても本質的な変化にはつながらないのではないかと——構造変化モデルで考えるとそうした視点に気づくことができます。

日本企業がデジタル化で顕著な成果を出せなかったのは、局所的な生産性向上や置き換え（人員削減等）にとどまって、ビジネスモデルの転換や産業構造の転換にまで踏み込まなかったからだと思います。既存の構造を前提とした発想では、差別化できず、超過利益を獲得できないのです。

## アーキテクチャの決定と進化のプロセス

- アーキテクチャ＝サブシステム間の役割分担とインターフェースの設計
- アーキテクチャは制約となっているプロセス(ボトルネックプロセス)を最大限有効活用するように決定される
- 技術はボトルネックを(1)緩和するか、(2)除去する
- 従来のボトルネックが解消された時に新たなボトルネックが浮上する。

では、アーキテクチャ(構造)とはそもそも何なのか。社会システムも技術システムも、複数のサブシステム(またはモジュール、コンポーネント)で構成されています。それらサブシステム間の役割分担とインターフェースを設計することがアーキテクチャだと言えます。

例えば、世の中を政府・企業・非営利部門のように分けるのも、三権分立(立法・行政・司法)も、東アジアの安全保障のために米軍と自衛隊でどう役割分担するかというのもアーキテクチャです。最近では、技術的ソリューション、法的ソリューション、インセンティブ構造などを組み合わせてネットワーク社会のアーキテクチャをどのように設計するかということを語る法学者の人たちもいます。

このように「アーキテクチャがどう変化するか」という視点・フレームワークを固定しておき、テクノロジーがそれをどう変化させるのかを考えるというのが有効なアプローチだと思います。

では、何がアーキテクチャを決定づけるのか——。多くの方は先進技術の最も優れた機能に注目しがちなのですが、長年研究していると構造を決めるのは「ボトルネック」の部分だということが分かってきたのです。

例えば、コンピューターのOS(オペレーティングシステム)。初期のコンピューターはハードウェアとソフトウェア一体で、計算のたびに配線(ワイヤリング)をやり直していました。ハードとソフトが分離したのは、この作業があまりに面倒で非効率だったからです。すると、ハードはムーアの法則で急速に進化しましたが、ソフトの生産性はそのペースに追いつけなかった。そこで、古いソフトを新しいハードで動かすための翻訳層としてOSが生まれてきたのです。IBM360を起点に、ソフトとハードの分離が産業構造に大きな影響を与えました。ここで注目すべきは、生産性が低く遅れている部分(ソフトウェア)が全体の構造を決めたということです。

サプライチェーンでも同様でした。1980年代の終わりから90年代の初めにかけて、大手小売企業がこぞって電子受発注システムを導入しました。ところが、競争優位を築いたのは電子受発注を徹底させた企業ではなく、物流システムの改革に取り組んだ企業でした。その象徴がウォルマートです。先進技術はどんな企業でも導入できましたが、最大のボトルネックだった物流の改革に取り組んだ企業はほとんどありま

せんでした。ウォルマートはそのボトルネックに挑むことで抜きんでることができたのです。この例でも、全体の足を引っ張っているところに注目すると、次に何が来るのかが見えてくるのです。

## AIは「責任」を取れない

**菊池** AIに関してはハイエンドの議論ばかりされていますが、社会全体、産業全体の構造を変える要因はむしろローエンドにあるという視点はとても新鮮です。

**國領** ここから先はほとんど推測に近い仮説ですが、以上を前提にするとAI社会においては「責任」というものが構造を決める重要要素になると思うのです。

### 「責任」が業界構造を決める？

仮説：AI活用の具体的な形を決める上で誰が責任を取るのかが決定要因となる

- AIはほとんど何でもできるが責任を取ることだけにはできない
- もし製造物責任原則を徹底するのであれば、サービスモデルが有力？エッジが使うデータも中央管理？
- それはつまり垂直統合の再来を意味している？

対抗：データのネットワーク外部性⇒水平分業強まる？

古代ギリシャでは責任を負うのが市民で、それ以外は奴隷

AIはいろいろなことができますが、「責任」を取ることはできません。なんとなれば、AIは死刑を恐れないからです。死刑を恐れない存在は究極の責任主体にはなれないのです。古代ギリシャでは責任を負うのは市民であり、奴隷は免責されていました。つまり、責任を負うか否かで市民と奴隷の間に線引きをしたのです。現代においても同様に、誰が物事の責任を取るのかが社会や産業の構造を決めるのです。

AIを使ったモノやサービスで問題が起きたときに誰が責任を取るのか——。この問いへの答え次第で



これからの業界構造が変わります。

例えば自動運転車の事故の扱いについては製造物責任の議論に収束しているようなのですが、自動運転を作ったメーカーや販売会社に全責任を負わせるのは難しいでしょう。なぜなら、自動運転はハード（クルマ）・ソフト（AI）・データ（情報）が一体的に運用されることで初めて成り立つものだからです。AI本体や運行データを持たない自動運転車メーカーに事故の全責任を負わせることはできません。他方、ハンドルを握ることもなくなってしまった利用者はこれまでのような運転者責任を負う気にはなれないでしょう。

だとすれば、自動運転車についてはメーカーがユーザーに売り切ってしまう「販売モデル」ではなく、ハード（クルマ）・ソフト（AI）・データ（情報）を垂直統合的に管理・運用・保守する「サービスモデル」でなければ結果責任を負うことができないのではないか、という仮説が浮上してきます。Google からスピンアウトした Waymo が提供する自動運転タクシーは自動運転をサービスモデルとして具現化したものです。もしかすると、AI 時代には「垂直統合」が産業構造の主流になるかもしれません。

## ビッグブラザー AI による寡占は考えにくい

ここで、「AI によって寡占化が進むか?」というテーマが浮上してきます。究極の仮説は、一つの巨大 AI（ビッグブラザー AI）が世界を支配するようになるということです。より多くのビッグデータを持っている方が有利なので自ずと寡占化が進むという考え方です。

一方、対抗仮説としてスモールデータのエージェント AI が林立し、AI 間の相互ネゴシエーションによって部分最適を図ることで全体が均衡するというシナリオもあり得ます。

例えば物流業界を考えると、あらゆるトラックとあらゆる荷物の状況を単一のビッグブラザー AI が把握

### AI によって寡占化が進む?

仮説: 一つの全知全能 AI が世界を支配するようになる

- より多くのデータをもった企業が勝っていく?
- そこまでいかずとも当面は寡占化が進む?

対抗: パーソナルエージェントや企業エージェントが林立してネゴシエーションをする世界になる

1970年代になされた市場経済 vs 計画経済議論の再来。  
当時はハイエクなどが主張した市場モデルが勝った

すれば究極の全体最適が実現するはずですが、実際にはそんなことになりません。やはり、個々の企業の AI エージェントが自社のために最適化を行い、他企業の AI エージェントとネゴシエーションすることで部分最適を測るというモデルに落ち着いていく。私はそう思います。

再び、「責任」という観点で考えてみましょう。

ビッグブラザー AI が全体最適解を導く世界は、計画経済的、中央集権的で、それ故に責任の所在が不明確です。ビッグブラザー AI は極めて優秀ですが責任は取れません。

しかも、株主、顧客、裁判所（刑法・民法）は「全体最適だから一部の過失・損失はやむを得ない」などということを決して認めません。市場経済、グローバル・サプライチェーン、インターネットなど歴史を振り返っても、ほとんどの社会システムは分散調整型モデルに落ち着いています。なぜなら、「責任」を取

れるのは企業、法人、個人、国家、つまり法的・社会的に罰することができる主体だけだからです。

つまり、全体を一つの賢いAIが最適化するのではなく、責任を負う主体（法人・個人）ごとにエージェントAIがあり、それぞれが自分の利益と責任の範囲で判断し、交渉し合うという世界の方が現実的だということが見えてきます。これは、経済学者フリードリヒ・ハイエクの市場経済モデルの世界観を踏襲するものです。

## 認知限界の突破がアーキテクチャ再構築を迫る

**菊池** 「ボトルネック」「責任」のほかに、AI時代のアーキテクチャを決定づける要因はありますか。

**國領** AIの支援によって人間がこれまでの「認知限界」を超えていくと、それに連動して世の中のすべての構造が変化していく可能性があると思います。

ノーベル経済学賞受賞者でもある社会学者ハーバート・サイモンは、「限定合理性」を唱えたことで知られています。伝統的経済学は人間を完全な情報と計算能力をもつ合理的選択主体と仮定しますが、実際の人間は認知能力や情報が制約される中で必ずしも最適な解ではなく「十分に満足できる解」を選んで意思決定しているという現実に則した学説です。

### 認知限界とアーキテクチャ

複雑なものを「準分解可能」な下位システムの連結したものとして取り扱うことで限界を突破する

Simon, H. A. (1981). The sciences of the artificial (2d ed.). MIT Press.  
稲葉元吉・吉原英樹訳. (1987). システムの科学 (新版 ed.). パーソナルメディア.

もし認知限界が人工物（機械もソフトも組織も政治体制も）の構造の決定要因だとするならば、AIアシストで人間の認知限界が大きく変わろうとしている今、世の中のすべての構造が変化してもおかしくない

簡単に言えば、人間には理解力の限界があり、社会システムの構造は人間が理解できる範囲で決定づけられるということです。物質は分子や原子で構成されていますが、人間はそんな微細なスケールで世の中をとらえていません。ある程度の大きさをもったまとまり（chunk）の集合体として世の中を理解しています。

サイモンは、「複雑なものを準分解可能な下位システムの連結したものとして取り扱うことで限界を突破する」と表現しました。だとすれば、AIが人間の認知能力を肩代わりすることで限界を超え、それに連動して世の中のあらゆる構造が変化することになるかもしれません。

繰り返しになりますが、AIが社会に与える影響を考えると、「生産性」や「置き換え」という視点

だけに頼っていると大局を見失います。「ボトルネック」「責任」「認知限界」といった要素の変化によって全体の構造がどのように変わるのかという視点からの分析・研究が必要だと思っています。

**菊池** 構造変化モデルによってAI社会の輪郭が少しずつ見えてくることになりますね。重厚なお話をいただきありがとうございました。

## 國領 二郎 / Jiro Kokuryo

慶應義塾大学 名誉教授

早稲田大学ビジネス・ファイナンス研究センター 研究院教授

共愛学園前橋国際大学デジタル共創研究センター センター長

1982年東京大学経済学部卒。日本電信電話公社入社。1992年ハーバード・ビジネス・スクール経営学博士。1993年慶應義塾大学大学院経営管理研究科助教授。2000年同教授。2003年同大学環境情報学部教授などを経て、2006年同大学総合政策学部教授に就任し、2009年総合政策学部長、また2005年から2009年までSFC研究所長、2013年から2021年5月まで慶應義塾常任理事を務める。2025年4月より慶應義塾大学名誉教授、共愛学園前橋国際大学デジタル共創センター長、早稲田大学ビジネス・ファイナンス研究センター研究院教授に就任。彼は公的活動にも積極的に取り組んでおり、ボランティア団体との連携も深め、地域社会の発展に貢献した功績より総務大臣表彰を受賞している。

インタビュー実施：2025年12月10日

AI ガバナンスを巡る論点 2025

## <最前線>企業の AI 利活用

OpenAI の生成 AI「ChatGPT」が 2022 年 11 月にリリースされてから 3 年が経過した。日本の企業・組織における導入・利活用はどこまで進んでいるのか、今後の課題は何か——。AI 利活用最前線の声から AI ガバナンスを巡る論点をあぶり出す。

### Frontrunners

東京ガス	社長リードで迅速導入、失敗重ね勝ち筋に迫る
三井不動産	普及促進から付加価値創出フェーズへ
MIXI	3 カ月で全社 AI 化、強い危機感で推進
東京海上日動火災保険	情と理、両面からのアプローチが大切
博報堂テクノロジーズ	面白がって、使ってみて、行動してみる
NTT データグループ	2027 年度には社員 20 万人を AI 活用実践人財に

### Opinions

なぜ日本の医療界で AI 活用は進まないのか

国立循環器病研究センター西村邦宏 部長

AI の「安全」と「安心」を切り分けて考える

AI セーフティ・インスティテュート (AISI) 村上明子 所長

### Interviewer

菊池 尚人

デジタル政策フォーラム 代表幹事代理

慶應義塾大学大学院 メディアデザイン研究科 特任教授

※本文および図表の一部を割愛して掲載している。本文中下線は Web 版の URL リンク箇所を示す  
全文はデジタル政策フォーラムの Web サイトへ  
<https://www.digitalpolicyforum.jp/>

## 東京ガス 社長リードで迅速導入、 失敗重ね勝ち筋に迫る

清水 精太 常務執行役員 CDO に聞く



### AI 利活用の現状

「ChatGPT が 2022 年 11 月に一般公開され、さっそく私も使ってみました。生成 AI の潜在的な“破壊力”に大変な衝撃を受けました。東京ガスは DX への感度が高い組織ですが、目標設定の前提条件が根底から変わると思いました。まず経営戦略があってシステム戦略が決まるという常識が逆転し、生成 AI を前提として諸々の戦略が規定される時代が来ると直感しました。経営陣に諮ると認識は全く同じ。そこからの動きは非常に速かった」（清水 常務執行役員）

### IT 活用、DX を積極的に推進してきた社風

- ・ 既存のレガシーシステムへの対応、「2025 年の崖」（経済産業省の DX レポート、2018 年）への対応など、全社課題として DX を早期から推進（東京ガス DX サイト）
- ・ 大規模な顧客管理システムの導入など、デジタル技術が競争力を左右するという認識を強く持つ企業文化

### 迅速な生成 AI 全社導入

- ・ 社長の強いリーダーシップによる迅速な経営判断
- ・ 2023 年 7 月には希望する社員が生成 AI チャットツールを使用できる環境を構築（社員の半数に当たる 2,500 名超が日常的に利用、2025 年 12 月時点）

### プラットフォーム環境の整備、活用推進

- ・ 生成 AI ベースのアプリを開発するためのプラットフォーム環境の整備（生成 AI を搭載した社内アプリを独自開発・利用開始～ RAG 技術搭載のチャットツールや One to One マーケティング自動実行ツールを本格導入～、2024 年 10 月 10 日）
- ・ 回答・ナレッジ活用、テキスト分析・構造化、暗黙知の形式知化、図面作成・検図自動化、ペルソナ活用、提案業務高度化といった、横展開性が高く効果が出やすいユースケースや、業界固有・特定業務などの領域特化分野を中心にアプリ開発を推進
- ・ これまで延べ 110 件のユースケースの PoC を実施し、うち 37 件が業務適用、更に 25 件が業務適用目前となっており、約 60 件のアプリを実運用レベルで構築（PoC 中を含む、2025 年 12 月時点）
- ・ 原料調達から営業・マーケティング・カスタマーサービスに至るバリューチェーン全体での AI 活用を推進

## 意識変革と業務変革を同時推進

- ・ 技術導入への抵抗は少ない。AI の便利さが明確で、社員の使用意欲は高い
- ・ OpenAI の Sora2 を使った動画（清水氏本人の顔と芸能人をコラージュし、武道館で歌って踊る様子）を作成し、経営会議で投影。人間の能力を超えた圧倒的なスピード、可能性、使わないことのリスクについて意識合わせ
- ・ 複雑化、属人化した業務プロセスの改革が、AI を前提とした DX 推進には不可欠。業務プロセス改革を率いるグループマネージャーにコミュニケーション能力の高い人材を当てるなど、人事面からのアプローチも試行中
- ・ 総延長約 6 万 km のガス・インフラの工事に年間 1000 億円規模を投資している。図面作成→工事実施→出来形図作成→検収→データ化という一連のプロセスに生成 AI を適用。図面を OCR で読み取り構造化データ化するなど、AI を軸にした業務プロセスへの転換を推進中

## 付加価値創造とビジネスモデル変革

- ・ トップライン向上： デジタルマーケティングでの大量施策実行、エンドツーエンドのデジタル化によるボトルネック解消、EC 事業への注力でチャネルのデジタル化を推進、等
- ・ ボトムライン改善： 業務効率化から BPO までの徹底した人件費削減、営業利益率を現在の 5-7% から 10% 超のエクセレントカンパニーレベルへ向上させる
- ・ 新ビジネスモデル： エネルギー供給企業からデジタルソリューション提供企業への進化、パーティカル SaaS（特定業種特化型）での競争優位性確立、環境・デジタル・レジリエンスの 3 つの付加価値を統合提供
- ・ カスタマーゼロ： 社内で様々な製品やサービスを先行的に試行し、外部顧客に販売していく
- ・ AI エージェントを含むユースケース拡大を見据え、2025 年 6 月に「AI 活用方針」をグループ全体へ共有し、各組織の自律的かつ柔軟な AI 活用の“さらなる一歩”を後押ししながら、活用・開発の乱立や投資の個別化を防ぐとともに、社会的責任を担保した AI 活用・開発環境を迅速に整備することで、AI のメリット最大化と全社最適の両立を図る

## AI 導入リーダーとしての心がけ・気づき

「AI 利活用のベストプラクティスやベンチマークというものは確立されていなかったもので、先進ユーザー企業 20 社くらいに直接お話を伺いに行きました。どうしたら現場に浸透し、業務にレバレッジが効き、成果を出せるのか、私たちの状況もお話しながら具体的にディスカッションさせていただいた。ベストプラクティスはベンダー頼みではなく自分たちで探し作るもの。綺麗な絵を描いてもその通りにはいかない。微に入り細に入り目を配りながら実践を重ねるしかありません」（清水 常務執行役員）

## 他社から聞かれること／他社に話すこと

- ・ データ・アーキテクチャをいかに整えるか、データ・ガバナンスをいかに効かせるか： 当社も含めレガシー企業のレガシーシステムにおいて、AI 利活用を推進する上での最重要課題

- ・どのようなリーダーシップで進めたのか： 経営の率先垂範、現場での丁寧なコミュニケーション

## AI 導入で日本企業の競争力は上がるか？

- ・優勝劣敗ははっきりしていくはず
- ・当社が AI 利活用を即座に始めたのは、早く、たくさん失敗することができるから。失敗による自然淘汰の結果として成功が残る

## ソブリン AI について

- ・日本国内で、日本の価値観を前提とする企業にとってソブリン AI は必要。グローバルビジネスの場合には必ずしもそうではない。両側面がある
- ・ただし、企業や業界が蓄積してきたデータで AI を個別にチューニングすることはできる
- ・米欧が先行している大規模言語モデルに、日本勢が今からキャッチアップすることは難しいだろう。水平方向で真正面からぶつかるのではなく、業界特化型・カテゴリーキラー型の AI 利活用を深化させ垂直方向の戦いに持ち込む方が日本にとって現実的な戦略ではないか。それに対する政策支援パッケージなら歓迎したい

## 清水 精太 / Seita Shimizu

東京ガス 常務執行役員、CDO、ソリューション共創本部長

1995 年東京ガス株式会社に入社し、ガス工事の設計及び施工管理、パイプライン等のインフラメンテナンスに関する新工法の開発に従事した後、総合企画部等においてエネルギー環境政策対応を担当。さらに国内外の事業投資や事業開発を経験した後、2022 年より総合企画部長として中期経営計画策定、事業ポートフォリオマネジメント等に関する各種施策の立案、ソリューション事業ブランド IGNITURE の立ち上げを推進。2025 年 4 月より CDO、新規事業開発担当役員に就任。

インタビュー実施：2025 年 12 月 8 日

## 三井不動産 普及促進から 付加価値創出フェーズへ

山根隆行 DX 本部 グループ長に聞く



### AI 利活用の現状

「生成 AI 利活用を全社に広げていくため、社内 85 部門すべてに対して「AI 推進リーダー」を一人ずつ出してほしいと依頼したところ、あっという間に約 150 名が集まりました。AI 利活用に対する現場の興味関心の高さを実感しました。DX 本部と現場が一緒に取り組むことで、血の通った業務変革につながっていきます」（山根隆行 DX グループ長）

### 生成 AI 導入のステップ： 助走から本格利用へ

- ・ 2023 年に AI チャットツール（Azure OpenAI ベース）を内製し、社員向けに提供。文書作成・校正・調査などの業務効率化から開始
- ・ 2025 年 10 月、ChatGPT Enterprise を全社員が利用できる環境を全社共通 AI 基盤として整備。「カスタム GPT」により、部門や業務ごとのルールやマニュアルを組み込んだ RAG (Retrieval-Augmented Generation) 拡張により社内専用アシスタントの作成・共有が可能に（利用開始から約 3 カ月で約 500 件のカスタム GPT を運用）
- ・ 社員一人ひとりの業務効率化（全体底上げ）と、業務ごとのプロセス変革と高度化のからアプローチ。業務時間 10% 以上の削減を目指す

### 施策と組織

- ・ AI チャットツール導入時に全社員に研修を実施し、ChatGPT 活用のための基本スキルを定着。ChatGPT Enterprise 導入時にも全社員対象に研修を実施
- ・ 2025 年 4 月、DX 本部内に AI とデータに特化した DX 四部を新設。DX 推進体制の中に生成 AI 活用とデータ活用の司令塔をビルトイン（[三井不動産グループ DX 白書 2025-2026](#)）
- ・ 2025 年 11 月、全社 85 部門すべてから約 150 名の AI 推進リーダーを選出。集中的な勉強会ののち、草の根的・自走的な利活用の現場リーダーとして活躍中

### 代表的ユースケース

#### 【三井ホーム】

- ・ 営業ナレッジを言語化し、組織全体で共有・再現できる「営業支援生成 AI」を構築。AI エージェントを通じて誰もが知見を引き出せる環境を整え、住宅提案の属人化を解消

- ・顧客データと営業ナレッジを組み合わせることで、顧客一人ひとりの特徴やニーズ、状況に応じたアプローチ方針の立案が可能に
- ・2026年度に本格運用開始見込み

### 【三井不動産レジデンシャルサービス】

- ・フロントマネージャー約 600 人に対して「マンション管理業務支援 AI」を構築・提供
- ・マンション管理組合の理事会や総会を支援するための資料作成や議事録作成の自動化などによって大幅な効率化を図る。フロントマネージャーの能力（担当案件数）の拡張、サービス品質の向上を目指す
- ・2026年度に本格運用開始見込み

### 「社長 AI エージェント」により経営陣自らが旗振り

- ・2025年12月に、「社長 AI エージェント」の全社利用を開始（[三井不動産 | 社長 AI エージェントなど独自 AI 開発と、全部門で 150 名の AI 推進リーダーを設置](#)）
- ・植田俊 社長の公開情報や経歴、過去の発信内容、個人的なエピソードなどを取り込み、社長の「ものの見方・考え方」を立体的に再現した AI エージェント。中期経営計画などの経営情報も反映
- ・社員が社長 AI エージェントと対話することを通して市場環境、全社戦略、経営方針についての理解を深めてもらう。「経営と現場をつなぐ新たな基盤」を目指す
- ・2025年10月には、DX 本部内で「DX 本部長 AI エージェント」を利用開始。本部長向け説明資料の事前レビューによって本部長意向との不整合による手戻りが減少し、DX 本部員の資料作成・修正にかかる時間を平均約 30%削減

## AI 導入リーダーとしての心がけ・気づき

「2026年度からは AI の投資対効果（ROI）をしっかりと見ていく段階に入ります。ただ、一方では、短期的な ROI だけに固執し過ぎないことも大切であると感じています。AI は非常に速いペースで進化・発展しているので、その流れにきちんと乗っていること、AI を業務プロセスの中にきっちり入れ込んでいくこと、AI の変化にしっかりと食らいついていくことが一番大事だと思っています」（山根隆行 DX グループ長）

### 部門ごとにサイロ化したデータの統合が課題

- ・事業部門ごとにデータがサイロ化している、AI で活用するのに十分なデータが整っていない、といった課題がある状況。生成 AI の利活用で当社ならではの付加価値を生み出していくためには、AI が必要としているデータを十分に活用できる環境の整備が必要になってくる
- ・新設された DX 四部で取り組みを開始している
- ・データ統合による顧客価値や利益創出の実効性、投資対効果の見極めが重要

### 普及促進から付加価値創出へ

- ・ 社員の月間利用率 80%を目指しつつ、エージェント型 AI による業務プロセス変革、利益創出を進めていく。2026 年度の焦点は「全社普及促進」から「付加価値創出」にシフトへ
- ・ 何らかの利益貢献 KPI を設定することになるだろうが、あまり固執しすぎると大胆な発想・挑戦にブレーキをかけることになりかねない。適切な KPI 設定が大切
- ・ 不動産デベロッパーのビジネスモデル（用地取得→開発→運営）を根幹から変革するということまで、少なくとも現時点では想定していない
- ・ エージェント型 AI の活用が本格化すると、ハルシネーションのリスク、著作権侵害のリスク、セキュリティリスクなどが顕在化してくると見られる。利活用促進と並行してガバナンス体制の整備・強化を進める
- ・ 文系出身者が多い組織だが生成 AI に関しては現場からの反応がとても良い。自然言語 UI の分かりやすさ、取っつきやすさが大きい。部門を超えた横の連携も活性化している。こうした現場コミュニティの力を最大限に活用することが収益化フェーズにおいても重要なカギとなる

## 山根 隆行 / Takayuki Yamane

三井不動産 DX 本部 DX 四部 DX グループ グループ長

KDDI、旭化成などを経て、2023 年に三井不動産へ入社。DX 人材育成制度の企画推進を担当した後、2025 年 4 月より現職。

インタビュー実施：2026 年 1 月 8 日

MIXI

## 3 カ月で全社 AI 化、 強い危機感で推進

村瀬 龍馬 取締役 上級執行役員に聞く



### AI 利活用の現状

「AI 利活用がなかなか進まないの、2025 年 3 月に役員と部室長以上を集めた研修を行いました。すると、使い方が分からない、部下に任せている、他社の対応状況を知りたい、まず PoC を回そう、経営の意思が伝わってこない、といった声ばかり……。そういう考え方をリセットしましょうと訴えました。前例は、他社は、経営がって言うのをいったんやめて、それぞれの部署を自分自身がリードして、現状を突破するための施策を考えましょうよ、と」（村瀬 取締役 上級執行役員）

### 生成 AI の登場から導入初期まで

- ChatGPT リリース直後、経営陣（社長・取締役）がその可能性に気づき利用を開始。AI ベースの経営に舵を切る
- 2023 年 4 月から「ChatGPT Plus」の利用料補助を福利厚生として運用開始、感度の高い社員が個人単位で使い始める
- 2023 年 11 月から自社開発した生成 AI ツール「Chat-M」（Azure OpenAI ベースの LLM サービス）を社員向けに提供開始
- 2024 年 4 月、AI や業務改善に関する相談窓口となる全社横断組織「DX 推進グループ」を設置
- 2024 年 6 月、生成 AI 利用ガイドラインを全面改訂
- しかし、全社浸透はなかなか進まなかった。経営だけがやる気でも会社は変わらない・変わらない

### 転機＝AI 推進委員会の発足、部室長以上向け集合研修の実施

#### 【AI イノベーション・戦略委員会（AI 推進委員会）】

- 2024 年 12 月発足。メンバーは経営陣および全部門の部室長（のちに各部署から選出された「AI アンバサダー」も参画）
- 全社 AI 戦略文書（2025-2026）の策定・実行
- AI エンジニア／利用者のレベル設定と給与連動
- 経営アジェンダから戦術実行までの一貫サポート
- 本部長・部室長の役割を明確化：リード（丸投げ禁止）、報告（毎月レポート提出）、推進（ROI 算出、施策提案、等）

## 【部室長以上向け集合研修】

- ・ 2025 年 3 月実施。AI 委員会メンバーを中心とする約 80 名を 2 日に分けて
- ・ 極力スマホ・パソコンを触らず丸一日箱詰め (9:00-18:00 + 懇親会)
- ・ 単なる AI 利用の研修にとどめず、講義の中心は、心理的安全性、Fireside chat (役員との直接対話)、部室長による各部署での活用方針・AI 変革宣言
- ・ PoC はやらない。業務に即適用し最速で成果を出す方針を明確化
- ・ 満足度は Day1 が 97%、Day2 が 100%

## 集合研修後、一気に現場浸透

- ・ 2025 年 4 月のテーマ： 全員が AI をベースにした考え方へ (社員用研修資料やポータルサイトでの展開)。アンバサダー制度 (部室長+自発的メンバー) で部署間の横連携を促進。気軽な雑談の場、知見共有の場を整備
- ・ 2025 年 5 月のテーマ： AI ベースのモノを作る (成果・結果報告の習慣化、可視化)
- ・ 2025 年 6 月： AI を触ってみるからの脱却 (ROI を出す)
- ・ 2025 年 7 月～： AI がないと機能しないワークフロー標準へ ([「AI がないと仕事が進まない」月 1 万 7600 時間の業務削減を見込む MIXI の AI 活用 | ビジネスジャーナル](#))
  - ― 見積もりは従来比 80 ~ 90%削減からスタートし、AI を使わないと間に合わない状態に
  - ― 自部門の利益追求は進展。AI を横串に刺して部門を越境した展開へ
  - ― AI による製品の差別化、圧倒的なスピード感
- ・ 全従業員 (約 2000 人) の 99.02% が日常で生成 AI を使用
- ・ 画像・動画作成、プロモーション素材作成などで劇的な時間・コスト削減効果 ([MIXI DESIGN の AI 実践を加速するために。生成 AI ガイドラインをアップデートした狙いについて | Cocoda](#))
- ・ ChatGPT Enterprise を全社導入 (2025 年 3 月) し、2000 個以上のツール・チャットボットを作成・共有。法務・知財関連の問い合わせ、マネジメントツール、ゲーム事業等のデータ管理、デザイナー向けプラグイン生成・管理、等々
- ・ バックオフィス業務での効果が絶大。知財、財務・会計、人事、経営企画などでの AI エージェント活用が進展
- ・ 2025 年度は AI 活用によるコスト削減を通じて、約 10 億円の利益改善。特に外注費の削減が大きい。労働時間は月間推定 17,600 時間を削減見込み

## 今後の計画と展望

- ・ 全社展開と収益ドリブンの徹底: 利用率 99.02% から 100% へ。売上・利益を直接伸ばす AI ユースケースを積極創出
- ・ 少数精鋭チーム (5 ~ 6 名) で新規事業を量産、グローバル市場に展開
- ・ 社内に取り残されているサイロ化データもすべて検索できるようにする。すべてのデータは AI のため、Agent が動きやすくするためのものへ
- ・ 3D 生成 AI をワークフローへ統合し、ゲーム開発のアセットコストを抜本低減
- ・ AI によるコーディングの推進、体制整備

- ・採用においては「生成 AI を利用したことがある」を前提に
- ・「AI を使いこなす社員が高報酬を得る」を原則とする（AI を使うと生産量が増えるので評価も上がる）

## AI 導入リーダーとしての心がけ・気づき

「AI 導入で仕事がなくなり人が減らされるという脅威論がありますが、真逆だと思っています。人を削るのではなく作る量を増やす。今の人数で 2 倍、3 倍どころか 10 倍のプロダクトを作る、爆速で。そして、日本の枠にとどまらず、グローバル市場で売りまくる。AI 時代に勝ち残るための厳しい競争が既に始まっているのです」（村瀬 取締役 上級執行役員）

### 人間が最大のボトルネック

- ・ AI によって一つひとつの業務が迅速化しても、「人間の合意形成」がボトルネックになってなかなか結論・決断に至らず全体のスピードが落ちる（1 週間後の会議で決めるなど論外）
- ・ 合意形成コストを最小化するためにはチーム単位をコンパクトにすること、権限移譲を徹底することが必要
- ・ リスク・コスト分析と選択肢抽出は AI に任せた方が速い。最終の決断は、リーダーシップに委ねるか、新たな合意形成ロジックを策定するか、そのどちらかしかないだろう
- ・ 取締役の役割は「監督」と「執行」。AI は監督が得意。ならば、取締役は「執行」に戻って、人間ならではの手触り感のあるモノづくりに集中できるようになる。経営陣の役割が変わる
- ・ 一方で、過去 1 年の変革を振り返ると「人の多様性」に心底驚く。技術の変化に対する人間の適応力は素晴らしい

### なお残る強烈な危機感

- ・ AI 環境の変化は速く、数週～1 カ月単位で世界が変わる。経営のスピードやモノづくりのスピードを AI に合わせる必要がある（まだ足りない）。漫然としていけば 2 年以内に置いていかれる
- ・ AI の力を借りれば少人数で良いモノを大量かつ高速に作れるようになる。そうなった時、社員が MIXI に留まる理由は何か、社員を引き寄せる MIXI の魅力は何かが問われることになる。経営陣として、そのことを噛み締めながら進まなければならない

### 村瀬 龍馬 / Tatsuma Murase

MIXI 取締役 上級執行役員

2005 年に株式会社イー・マーキュリー（現 株式会社 MIXI）に入社。SNS「mixi」の開発に携わる。2009 年に退職後、ゲーム会社などを経て 2013 年に再入社。主にモンスターストライクの開発業務に従事。2016 年 7 月、XFLAG スタジオ ゲーム開発室長に就任し、XFLAG のエンジニア全体を統括。2018 年 4 月、執行役員 CTO 就任。2019 年 6 月、取締役執行役員 CTO 就任。2023 年 4 月より現職。

インタビュー実施：2025 年 12 月 25 日

## 東京海上日動火災保険 情と理、 両面からのアプローチが大切

牧原卓也 CX・dX 推進部長に聞く



### AI 利活用の現状

「DX も AI 利活用も、情と理の両面からのアプローチが必要だと思います。この会社を変革していこう、会社をもっと良くしていこうと現場で話すと、皆さん、とても熱く語り合い始めるんです。そういう情の部分が変革の推進力です。そして、ありがたい将来像を描き、現在地を確かめ、バックキャストして道筋を知り、KPI を設定して PDCA を回すという理の部分がその上に乗ってくる。まだ道半ばですが、当社の AI 利活用はそんなふうにして動き始めています」（牧原 CX・dX 推進部長）

### AI 利活用の 3 つの目的

- ・ 生産性向上によるリーンな業務運営
- ・ データを活用したアンダーライティング（リスク評価と保険条件の設定）の効率化・高度化
- ・ 顧客接点の拡大と価値提供の向上

### 生産性向上によるリーンな業務運営

- ・ 代理店からの照会応答に関し、電話からメールへの移行を推進中。移行の取り組みにより蓄積されたデータは全国で約 100 万件。これらのデータを活用し、生成 AI を業務フローに組み込み、事例を選択したうえで回答案を生成する仕組みを構築（複雑性の高い保険領域に特化した照会応答システム「AI Search Pro」を共同開発～照会応答時間を大幅削減し、お客様への新たな価値提供の拡大へ～）
- ・ 2024 年 11 月にサービスインし、現在、照会全体の 4 割で活用されている。これにより、電話入電量は約 3 割削減した
- ・ 代理店支援を目的に、代理店システムに AI を組み込むための PoC が進行中

### 社内版 ChatGPT の展開

- ・ 2023 年 10 月、セキュアな環境を構築し、全社員向けに提供開始（全社員向け生成 AI “One-AI for Tokio Marine” の活用開始～ ChatGPT による業務効率化を実現～）
- ・ 生成 AI 初心者向けに「クイックスターター機能」を実装。利用目的別にワンタッチでプロンプトが入力される仕組み。プロンプト事例集の配布や研修も実施中
- ・ 部門ごとの業務マニュアルを読み込ませ、部門のデータを基に AI 検索が可能となる社内 RAG（Retrieval-Augmented Generation：検索拡張生成）機能を提供

- ・現在の MAU（マンスリー・アクティブユーザー）は約 65%（パート社員を除く全社ベース）、月 10 回以上利用のミドル・ヘビーユーザーは約 41%と年度初から倍増
- ・その他、Copilot 無償版、ChatGPT Enterprise（ヘビーユーザー向けに約 2500ID を配布）なども活用（[生成 AI を活用した業務基盤進化に向けた OpenAI との戦略的連携について ～付加価値の高い顧客体験を実現する AI エージェントの実現に向けて～](#)）

## データを活用したアンダーライティングの効率化・高度化

- ・ドライブレコーダーデータの活用（[東京海上日動のドライブレコーダー付き自動車保険「ドライブエージェント パーソナル」](#)）
- ・課題は、事故時に取り付ける見積書の PDF データをはじめとする非構造化データが多いこと。これらを含めたデータ基盤の整備を推進中
- ・海外の保険会社は AI によるアンダーライティングの高度化で先行している。ベンチマークとして動向をしっかりと把握していく

## 顧客接点の拡大と価値提供の向上

- ・中期経営計画 [Re-New2026（2024～26 年度）](#) で、「本当に信頼されるお客様起点の会社」「リスクソリューション（保険 + α）で次代を支える会社」の 2 つを掲げた。顧客との接点をもっと増やしていくために AI を最大限活用する
- ・2023 年より取り組んでいる NPS（ネット・プロモーター・スコア）のデータをもとに顧客からの評価データや既存契約データを統合・活用することで顧客理解の深化を図るカスタマー・データ・プラットフォーム（CDP）を 2026 年 2 月にリリースする。また、同 CDP による顧客理解を基盤として、将来的には AI を活用したパーソナライズされた提案につなげる仕組みを検討する
- ・コンタクトセンターへの AI 導入を 2026 年 3 月目途に実施。ボイスボットによる初期対応の自動化、Speech to Text と生成 AI を組み合わせたオペレーターへの回答案サジェストや対応記録作成の自動化によるアフターコールワークの削減を企図。ウェブ（月間約 150 万 UU）との連携も強化し、問い合わせ対応の自動化比率を高めていく

## 人材育成

### 4 階層に分け、①役割と②育成方針を設計

DX リーダー（経営層・部店長）

- ① DX 戦略の主導と組織風土の醸成
- ② 知識と意識の両面の習得をめざし、部店長向けに早朝講座やワークショップなどを開講

DX コア人材（本社）

- ① データおよびデジタルに関する施策の立案と業務高度化
- ② スキルセットを把握し、育成計画に反映

DX ドライバー

- ① 各組織の DX 推進リーダー
- ② 伴走支援を通じた、DX リテラシー・スキルの習得および組織推進強化

全社員

- ① DX リテラシー・スキルの習得および社内デジタルツールの徹底活用
- ② 各業務での活用、動画コンテンツの活用促進、IT パスポートの資格取得推奨

※従来の「DX 人材」の目指す姿について見直しを考えている。「AI に的確な指示を出し、業務で使いこなす」という能力が必要。それに応じて育成方針も変えていく

## AI 導入リーダーとしての心がけ・気づき

「AI 利活用の推進役としてもどかしく感じているのは、現場のみなさんが AI の効果を実感して心から喜んでいるかということ、まだそこまでには至っていないことです。目に見える効果を生み出していくこと、効果を実感できる分かりやすい指標を設定すること、この2つが最大の悩み。ただし、根っこのところはお客様に価値提供するというところに尽きる。現場の思いもそこにある。肝に銘じておきたい」（牧原 CX・dX 推進部長）

### 推進上の課題とアプローチ

- ・現場において導入効果を実感できるような評価指標の設定は課題。例えば、代理店さんからの照会応答に関する入電が半減したとしても、必ずしも実感を伴った業務量の減少に繋がっていないと感じている
- ・AI 活用による将来像をナラティブに未来像として語りつつ、「今」の数字で課題を認識してもらい、実際の行動を促す。こうしたビジョンと KPI の往復でムーブメントを創出することが求められる
- ・人事制度、営業体制、代理店システムの改革など、様々なプロジェクトを横断し、全社視点で AI 施策をビルトインすることが必要
- ・全社を俯瞰する経験と視点、社内人脈を駆使した支援者拡大がカギ

### 牧原 卓也 / Takuya Makihara

東京海上日動火災保険 CX・dX 推進部長

1996 年、東京海上火災保険株式会社入社。

京都中央支店、経済産業省出向、経営企画、労働組合専従、東京新都心支店、滋賀支店（支社長）、あんしん生命出向（営業企画・デジタル戦略）を経て、2024 年より現職。

インタビュー実施：2025 年 12 月 23 日

## 博報堂テクノロジーズ 面白がって、使ってみて、 行動してみる

木下陽介 執行役員に聞く



### AI 利活用の現状

「生成 AI によって我々のクリエイティブワークが奪われてしまう、というような心配はしていません。むしろ、クリエイティビティを拡張することにつながる。AI との付き合い方は、想定するリスクに配慮しながら、面白がって、まず使ってみて、行動してみる、ということに尽きるのではないかと。博報堂DYグループは元々そういうカルチャーの会社です。当初は少し戸惑いもありましたが、「with AI」の業務スタイルが一気に浸透していきました」（木下 執行役員）

### 統合マーケティングプラットフォーム「CREATIVITY ENGINE BLOOM」の開発

- ・ 木下執行役員が開発・推進のリーダー
- ・ トップクリエイター TBWA HAKUHODO の細田高広チーフ・クリエイティブ・オフィサー（CCO）のコンセプト開発手法を AI で再現した STRATEGY BLOOM CONCEPT、通称「細田 AI」を全社展開。得意先のマーケティング業務／事業成長を支援（[AI と人が共創するコンセプト開発 ～ CREATIVITY ENGINE BLOOM Vol.2 | グループトピックス | 博報堂DYホールディングス](#)）
- ・ 細田 AI の利用社員は約 2000 名に達し、コンセプトワーク労働時間を約 4000 時間削減

### イマジネーションパートナー「バーチャル生活者」の開発

- ・ 2024 年 3 月にプロトタイプを発表
- ・ 2025 年 11 月に“エビデンスベースド”「バーチャル生活者」をリリースし全社展開。博報堂が保有する豊富な生活者データを基に AI によって様々なペルソナの生活者を再現。バーチャル生活者との対話によって、マーケティング施策の立案に活用推進中（[博報堂DYグループ、AI がデータに基づき再現した複数の生活者との会話を実現生活者発想を支援する”エビデンスベースド”「バーチャル生活者」を開発 | コーポレートニュース | 博報堂DYホールディングス](#)）

### AI 利活用方針「Human-Centered AI Professionals」

- ・ AI 関連の人財や知見、データ、ネットワークなどのリソースを結集し、企業のマーケティングや事業成長を総合的に支援する専門集団として組成
- ・ AI 活用は単なる DX ではなく、マーケティングのあり方そのものを進化させる基盤と位置づけ（[Human-Centered AI Professionals - 博報堂DYホールディングス](#)）

- ・ AI による新しいマーケティングでクライアントのビジネスを支援する
- ・ 目指すのは、マーケティング業務プロセスの「効率化」と「高度化」

## 導入と推進

- ・ トップダウンの意思決定で Gemini を博報堂全社導入
- ・ 早期に AI ガイドラインを整備し、安全に利用可能なルールを策定
- ・ AI ガイドラインの枠内で「まず使ってみる」という行動原則を推奨
- ・ 当初は AI 活用に対する戸惑いも一部にあったが、成果が共有されるにつれて「with AI」の業務スタイルが急速に浸透していった
- ・ 「人間中心の AI」観をベースに、AI とともに新たな問い、発想、未来を探求するという利活用方針を打ち出す。「AI と正解を出す」のではなく「AI と別解を出す」ことを目指す

## ワークショップ活用によるクライアント共創

- ・ クライアントのニーズに合わせた Gems（ユーザーごとにカスタマイズできる Gemini の機能）を HCAI Professionals のメンバーにより、共創型で業務を進める取り組みを拡大
- ・ クライアントとのワークショップでの AI 活用を推進。コンセプトワーク、アクションプラン、アイデアの生成・洗練化、ドキュメント作成、未来シナリオ作成など。プロンプト設計にこれまでの経験知が活かされている。AI を活用したクリエイティブワークのテストベッドになっている
- ・ 顧客から「社員がなかなか使ってくれない」という相談を受けることもある。生成 AI 普及のためのエバンジェリストを選任して普及啓発するコンサルティング・メニューは意外に好評。AI 活用に関心する相談も増えている

## 組織・働き方の変化

### 【ポジティブな変化】

- ・ プロジェクト初期にかかる時間が短縮、クリエイターは考えることに時間を多くとれるようになった
- ・ マーケ職がクリエイティブ、クリエイティブ職がマーケ業務など、越境による多技能化が進展
- ・ 働き方改革の推進（長時間労働の抑制）、DX 受容度がアップ

### 【ケアすべき要素】

- ・ 過度に「正解」を求める傾向、「即納品可能な答え」を求める傾向が初期にはあった
- ・ 失望感が高まると利用停滞のリスクがあるため期待値コントロールが必要
- ・ 「正解」ではなく「別解」を引き出すものであるという考え方の浸透

## 効果測定の考え方

- ・ 「労働時間の短縮」と「成果物の質的向上」の両面からの測定・評価が必要
- ・ 単なる「利用人数がどれだけ増えたか」を成果指標とはしない
- ・ 広告会社で以前よく行われていた「広告コピーを 100 本書く」といったタスクは AI なら数秒でアウトプットできてしまう。AI に任せられることは任せ、人間は人間でなければできないことに集中し、全体の付加価値を上げていきたい

- ・ 現在、より短い時間でより高品質な成果につながるような指標設計に取り組んでいる

## AI 導入推進活動における心がけ・気づき

「今、マーケティングを取り巻く環境が大きく変わりつつあります。テレビ広告がダイレクトマーケティングに結びついたり、リターゲティングが中心だったデジタルマーケティング領域でも潜在顧客への新たなアプローチが求められたりと、従来の定石を超えた高度なマーケティングが求められるようになりました。この混沌とした状況を突破するために、AI の力を借りた新しいマーケティング手法を開発し、クライアントの期待に応えることが我々の最重要課題です。そこで必要になってくるのは、この変化を楽しみ、好機到来、働き方も暮らし方ももっと良くなると解釈できるポジティブなマインドセットだと思います」（木下執行役員）

### 人材育成・能力開発の方向感

- ・ 生成 AI を使って簡単にコーディングができるバンプコーディングやノーコードまたはローコードで生成 AI アプリケーションを開発・運用できるオープンソースプラットフォームが整備されてきている。クライアント向けの AI エージェントを担当コンサルタントが作成するようなことも可能になってくる。AI エージェントの開発や生成 AI の便利な使い方などの AI 活用ナレッジを、社員一人ひとりが「ベーススキル」として習得していくことが必要になってくるのではないかと
- ・ それ以上に重要なのが「ソフトスキル」。様々な専門性・価値観を持つメンバー間で、AI による出力を基にして付加価値の高い結論を導くための「コミュニケーション能力」「会議体・ワークショップ運営能力」「課題解決に向けた協働能力」などが求められるようになる。そうした能力開発手法の設計に着手したい

### 木下 陽介 / Yosuke Kinoshita

株式会社博報堂テクノロジーズ 執行役員、マーケティング事業推進センター

株式会社博報堂にて、マーケティングやコンサルタント職として多数の業種の企業を担当。2010 年より研究開発職としてマーテック、アドテック、AI や XR 技術を活用したマーケプロダクト開発を推進。現在、博報堂 DY グループの統合マーケティング基盤「CREATIVITY ENGINE BLOOM」開発のリーダーを務める。

インタビュー実施：2025 年 12 月 22 日

# NTT データグループ 2027 年度には社員 20 万人を AI 活用実践人財に

正野勇嗣 Head of Global AI Office（室長）に聞く



## AI 利活用の現状

「システム開発の現場では AI は、使うか使わないかではなく、どう使うかのフェーズに入っています。AI 活用が主流化し、SIer が提供すべき価値も変わりつつあります。人・月リソースの提供という労働集約型の発想から抜け出さなければなりません」（正野勇嗣 室長）

### システム開発 / SIer 領域に変化の認識

- 生成 AI は「使うか、使わないか」を議論する段階を過ぎ、「どう使うか、どこまで使いきれるか」のフェーズに移行
- コード生成など単体タスクの自動化（レベル 1）から「複数タスクを統合した開発ロールの意思決定」（レベル 2）、「プロジェクト単位の AI が連動した開発シナリオの一貫した支援」（レベル 3）へと全工程に拡大
- これまでに、「特化型 AI」から「生成 AI」「AI エージェント」の活用へと進んできた。今後は、自ら目的を理解し計画・行動・学習を繰り返す「Agentic AI」、人間並みの知能を持ち幅広い課題に対応できる「AGI（汎用人工知能）」、人間を超える知能で社会や科学を設計できる「ASI（超知能 AI）」へと急速に進化していく
- SIer が提供すべき価値が大きく変わる。人的リソース提供の価値は消失し、「人間が作る」から、「AI が作る成果を保証しビジネスに変換する」役割へ
- 生成 AI 時代における SIer の 3 つの役割
  - (A) 変革パートナー： AI を活用した高速な業務改革・価値創出をリード
  - (B) AI-Native SIer： 開発プロセスを AI を前提に再設計し、品質と説明責任を担保
  - (C) 特定領域のディープスペシャリスト： 特定業界・技術領域に深く特化し、競争優位を確立
- 従来の「人・月を積み上げる SI モデル」は構造的に変化。製造工程に人・月が集中する「山型カーブ」から、前工程のコンサルティングの充実や後工程の品質管理や継続的改善に重点を置く「逆 U 字型カーブ」にシフトする

### NTT データの AI 活用戦略

- AI を最大限に活用し、「労働集約型」から「AI 駆動型ビジネスモデル」への展開を推進
- 「積極的な AI 活用の推進」と「AI ガバナンスの徹底」の両輪で取り組み、ビジネス拡大を図る

- ・「AIドリブン開発」の体制を早急に整備へ（[NTT データ、AI がシステム開発 IT 人材不足を解消 – 日本経済新聞](#)）

## 推進組織の整備

- ・ NTT データグループ（持株会社）に「Global AI Office」を設置（2023年10月）。グループ全体の AI 戦略の立案と投資を実行。また、グローバルでの法規制対応、コンプライアンス対応、リスク管理などを担当する「AI ガバナンス室」も併設している（2023年4月）
- ・ NTT データに「GenAI ビジネス推進部」を設置（2024年4月）。国内 AI ビジネスの加速を担当
- ・ 米 NTT DATA Inc. の Chief AI Officer（Abhijit Dubey CEO が兼任）が海外リージョンの AI 戦略立案と実行管理を管掌
- ・ 新会社 NTT DATA AIVista を米シリコンバレーに設立（2025年12月）。先端 AI 技術に基づく新たな価値創出、顧客の経営課題解決、業務プロセス変革の実現などを推進する（[AI ネイティブな新ビジネス創出を推進する新会社 CEO に Bratin Saha が就任 | NTT データグループ – NTT DATA GROUP](#)）

## AI 導入リーダーとしての心がけ・気づき

「変革パートナーとしてお客様の成果創出をお手伝いしていくための最重要課題は AI 人材の育成です。スキルの底上げとマインドセットの転換を急いでいます。2027年度には全社員 20 万人が AI 活用実践人材として世界中で活躍する体制を整えていきたい」（正野勇嗣 室長）

## 今後の課題

### 【People】

- ・ 経営層やミドルマネージャーのマインド変革： 既存業務の改善ではなく、業務プロセスそのものを AI 前提で再構築する変革意識が重要
- ・ 新ロールに対応した人材不足： 品質保証やビジネス連携が重要。顧客の最前線で AI を活用し実装していくエンジニアの育成が急務

### 【Process】

- ・ AI ネイティブ開発への移行： AI の恩恵を最大限に享受するためには、従来プロセスの延長ではなく、AI を中心としたプロセスの再設計が重要
- ・ ガバナンスと法的枠組みの整備： 全社統一の AI 利用ルール・ガバナンス策定に加え、倫理・法的枠組みの確立が必要

### 【Technology】

- ・ 技術の目利き力： 進化が速い AI 技術に対してグローバルでアライアンスを組むなど、インテリジェンス機能を設けて変化に追随。適性を見極めシステムに組み込んだ上で ROI を向上させることが重要
- ・ レガシーシステムとの接続や AI-Ready なデータ基盤の整備： AI 活用を阻む既存のレガシーシステムや構造化されたデータ基盤の整備が必須

## 「人財育成」が最重要課題

- ・ 全世界の社員約 20 万人を対象とし、生成 AI の人財育成フレームワークを策定
- ・ 昇段制度（Whitebelt → Yellowbelt → Greenbelt → Blackbelt）を設け、競争意識・成長実感を埋め込む
- ・ 生成 AI 実践人財（Yellowbelt・Greenbelt・Blackbelt）は 2025 年 10 月に 7 万人に到達。2027 年度には全社員 20 万人を目指す

## これまでの成功要因

- ・ 初期段階から全社員が利用できる環境を整えたこと
- ・ **Client Zero** として社内利活用を推奨・指針としたこと（社内活用と顧客提供の両輪）
- ・ 自社完結にこだわらず、社外との連携、アライアンスを積極活用したこと

## 正野 勇嗣 / Yuji Shono

NTT データグループ Head of Global AI Office（室長）

NTT データグループ全体の生成 AI 戦略および普及推進を担当。ソフトウェア開発技術の R&D に従事後、大規模開発プロジェクト支援、全社技術投資戦略の立案、グローバル Data&Intelligence 戦略・連携などに従事。開発自動化技術や AI 技術に関する雑誌・記事を執筆。

インタビュー実施：2026 年 1 月 9 日

## オピニオン

# なぜ日本の医療界で AI 活用は進まないのか

国立循環器病研究センター 西村邦宏 部長



日本の医療分野において、AI 活用の可能性は決して低くない。音声データによる認知症の早期検知、生活データを用いた高齢者の見守り、歩行情報や健診データからの疾病リスク予測など、技術的にはすでに実装可能な段階にある。実際、適切に設計された AI 活用によって、数千万円の投資で数億円規模の医療費削減につながった事例も存在する。

それでも、日本の医療界全体として AI 活用が本格的に進んでいるとは言い難い。その理由は、技術の未熟さではない。問題はむしろ、制度、投資、文化が絡み合った構造的な停滞にある。

## 「数が力」なのに・・・3 桁遅れの日本の医療データ

AI 活用において最も基本的な条件は、十分な量のデータが集積されていることだ。海外、例えば英国では国民規模で電子カルテや医療データが一元的に管理され、数百万人から数千万人単位のデータを用いた AI 活用型研究や機械学習モデルの開発が進んでいる。

一方、日本では医療データが制度ごと、事業ごとに分断されている。さらに、厳格な個人情報保護規制や運用ルールによって、横断的な利活用が極めて難しい。がん全体の1割を占める希少がんなど希少疾患などは、県人口が少ない県では個別疾患の集計が公開できないなど「患者数が少ないため分析できない」という本末転倒な状況すら生まれている。

彼我の差は3桁レベルである。AI にとって「数が力」である以上、データ利活用の障壁のため、日本の医療界は欧米、およびシンガポール、台湾などに比べて AI 活用のスタートラインにも立てていないと言っても過言ではない。

もう一つの大きな要因は、国家としての投資規模の違いである。海外では医療 DX(デジタルトランスフォーメーション) を社会インフラと位置づけ、兆円単位の予算を継続的に投じてきた。一方、日本の医療 DX 投資はここでも桁が二つ、三つ小さい。

標準電子カルテの導入が必要だと分かっているが、行政も努力は重ねているが全体の予算規模はまだ小さく費用を誰が負担するのかという問いに対して答えを導き出せない。結果として、医療機関も IT ベンダーも動けず、構想だけが宙に浮く。医療 AI は「民間の工夫・努力」に任せるにはあまりにも荷が重過ぎる。本来は国家戦略として政府が腹をくくって投資すべき領域だ。

## 情緒論が数量的議論を押し流す

日本の医療を語る際、避けて通れないのが情緒論の強さである。医療倫理の議論は、「一人でも不利益を被る人がいれば許されない」という情緒的な極論に傾きがちだ。その結果、AI 導入によって「何人を救えるのか」「どれだけ医療費を抑制できるのか」といったデータに基づいた統計的・数量的な議論が後回しになってしまう。また個人情報保護が欧米では公衆衛生分野で Public Good のため例外とされる

のに対して、日本では行政、民間企業、アカデミアもあまりにも慎重で石橋を叩き過ぎて割ってしまう状況かと思う。

政府が施策を打っても、事後検証や効果測定が十分に行われているとは言い難いので次につながらず一発の打ち上げ花火で終わってしまう。「やってみて、どれだけの成果が出たのか」を数字で語る文化・意識が弱いことが、AI 活用への社会的合意形成を難しくしている。

## 完璧主義が試行錯誤を許さない

医療現場の高齢化も、AI 利活用の大きな壁だ。電子カルテや AI システムの導入には莫大な初期投資が必要であり、引退が視野に入る高齢の開業医にとって、その投資は現実的な選択肢になりにくい。そうして、現状が維持されていく。若い世代ほどデジタルへの抵抗感は小さい。しかし、意思決定権を持つ層が上の世代に偏っているため、現場の感覚と経営判断の間にズレが生じる。この世代間ギャップが、変革のスピードを著しく鈍らせている。

日本社会には、「失敗してはいけない」という怖れの空気がある。医療という分野ではなおさらで、失敗は決して許されない。人々の健康や生命を守る仕事なのだから当然だが、度が過ぎると数字とエビデンスに基づく冷静な議論さえできなくなってしまう。AI を社会実装するためには試行錯誤が不可欠である。わずかなミスも許されないという完璧主義が、新しい取り組みへの挑戦を萎縮させてしまう。

## AI を使わなければ医療がもたない

ただ、状況は静かに、しかし確実に変わりつつある。医療現場の人手不足は深刻化し、従来のやり方を続けるだけでは立ち行かなくなりつつある。例えば、生成 AI の活用によって看護師の残業をゼロにした病院が出現したりと、導入の効果が明らかになり始めている。もはや論点は「AI を使うかどうか」ではない。人口減少による働き手の減少に伴い「使わなければ医療がもたない」という現実が、すぐそこまで来ている。

必要なのは、小さくてもよいから成功事例を積み重ね、その効果を数字で示すこと、そして公共の利益という視点から医療データの活用を社会全体で再設計することだ。医療 AI は魔法ではない。しかし、冷静な数理と現実的な投資判断があれば、日本の医療を持続可能にする有力な道具になり得る。その判断が、いま私たちに突きつけられている。 (談)

### 西村 邦宏 / Kunihiro Nishimura

国立循環器病研究センター 予防医学・疫学情報部 部長、医学統計研究部 部長  
神戸大学医学部 AI・デジタルヘルス学科 特命教授  
京都大学医学部卒。

2002 ハーバード大学公衆衛生大学院修士課程修了 (医療統計・疫学修士)

2003 ハーバード大学公衆衛生大学院修士課程修了 (医療経済学修士)

2005 神戸大学医学部立証検査医学講座助教 2007 年 同准教授

2018 国立循環器病研究センター予防医学情報部部長

2022 神戸大学大学院医学研究科 AI・デジタルヘルス学科特命教授 (兼任)

国立循環器病研究センター医学統計研究部 部長

インタビュー実施：2025 年 11 月 26 日

## オピニオン

# AIの「安全」と「安心」を切り分けて考える

AI セーフティ・インスティテュート (AISI) 村上 明子 所長



生成 AI の急速な普及により、「AI は安全なのか」「AI を信頼してよいのか」という問いが、これまでになく社会の前面に出てきました。私自身、AI セーフティ・インスティテュート (AISI、読みはエイシー) の活動を通じて、国内外でこうした問いと向き合ってきましたが、整理しておきたいのは、AI における「安全」と「安心」は同じものではないという点です。

### 「安全」は技術で作れるが、「安心」は技術だけでは作れない

私が考える「安全」とは、AI を使うことで客観的に被害が生じない状態を指します。モデル評価、レッドチーミング、ガードレール設計、サイバーセキュリティ対策などは、いずれも安全性を高めるための技術的な取り組みです。安全は、一定程度までであれば、技術と検証によって客観的に評価することができます。

一方で、「安心」は技術だけでは完結しません。安心とは、利用者や社会が主観的に「納得して使える」と感じられる状態です。たとえ安全性が技術的に確認されていたとしても、「よく分からない」「誰が責任を持っているのか見えない」と利用者を感じた場合、安心は生まれません。

ここで重要なのは、安心は結果ではなく、プロセスへの信頼から生まれるという点です。どのような考え方で設計され、どのような前提で使われ、問題が起きたときに誰がどう対応するのか。そうした姿勢や説明があって初めて、社会は AI を受け入れることができます。

### 日本では「安心」が先に問われる

日本では、特にこの「安心」が強く問われる傾向があります。海外では、まず安全性を最優先し、使いながら改善していくという考え方が一般的ですが、日本では「本当に大丈夫なのか」「誰か困る人はいないのか」という安心の問いが、しばしば先に立ちます。これは単なる慎重さではなく、日本社会が培ってきた価値観の表れでもあります。

ただし、この整理を怠ると、**安全に関する技術的議論が、安心をめぐる情緒的議論に飲み込まれてしまいます**。その結果、何が問題で、何を議論すべきなのかが曖昧になり、建設的な議論が難しくなってしまいます。

この文脈で、AI ガバナンスの役割を考える必要があります。AI ガバナンスは、「AI を完全に安全にする仕組み」ではありません。むしろ、**安全を、AI 利活用の前提としつつ、社会に安心をもたらすための判断と説明の枠組み**だと私は考えています。

重要なのは、AI ガバナンスが政府や専門機関だけの仕事ではないという点です。実際に AI を使い、その結果に対して説明責任を負うのは、企業や民間事業者です。したがって、組織レベルの AI ガバナンスは、国レベルの枠組みとは別に、主体的に設計されなければなりません。

## ソブリンAIではなく、「AIソブリンシー」という考え方を

ここで誤解されがちなのが、「ソブリンAI」という考え方です。「国産であれば安心」という発想は分かりやすい一方で、現実を単純化しすぎています。どこで作られたかよりも重要なのは、誰がそのAIを管理し、制御し、説明できるのかという点です。私は、重要なのは「ソブリンAI」ではなく、どこに主権的な判断を置くのかという意味での「AIソブリンシー」だと考えています。技術の国籍ではなく、統治の所在こそが、安心につながります。

AIの進化は速く、不安を煽る言説も少なくありません。しかし、恐怖に引きずられて立ち止まることも、無警戒に突き進むことも、どちらも望ましい姿ではありません。安全と安心を丁寧に切り分け、技術と社会の双方から備えていく。そのプロセスを誰が、どの立場で引き受けるのかを明確にすることこそが、AIを持続的に活用するための前提条件なのだと思います。 (談)

### 村上 明子 / Akiko Murakami

AI セーフティ・インスティテュート (AISI) 所長

1999年4月 日本アイ・ビー・エム株式会社 東京基礎研究所 入社

2016年1月 同社 東京ソフトウェア開発研究所

2021年4月 損害保険ジャパン株式会社 入社 執行役員待遇 DX 推進部 特命部長

2021年10月 同社 執行役員待遇 DX 推進部長

2022年4月 同社 執行役員 CDO(Chief Digital Officer) DX 推進部長

2024年2月 AI Safety Institute 所長 [現職]

2024年4月 損害保険ジャパン株式会社 執行役員 CDaO(Chief Data Officer) データドリブン経営推進部長 [現職]

2025年4月 SOMPO ホールディングス株式会社 執行役員常務 グループ Chief Data Officer [現職]

インタビュー実施：2025年12月17日



AI ガバナンスを巡る論点 2025

## <特集>AI×クリエイター

生成 AI の登場と進化は、映像クリエイティブにどのような影響を及ぼすのか——。若手のクリエイター、放送関係者、起業家、研究者たちの眼に、今、何が見えているのか？

### Creators

山口ヒロキ（映画監督）	抑圧された創造性の解放が映画を変える
篠田貴之（テレビ）	現場の現場による現場のための AI 導入
sola（15 歳 AI ネイティブ）	空気のように、感じるままに

### Viewpoints

松原冬樹（起業家）	AI × 安全保障において何が重要か
高木志郎（起業家）	研究という営みが根本的に変わりつつある
花光宣尚（研究者）	誰もが何でも作れる時代の創造性とは？

### Interviewer

菊池 尚人

デジタル政策フォーラム 代表幹事代理  
慶應義塾大学大学院 メディアデザイン研究科 特任教授

# 抑圧された創造性の解放が 映画を変える

映画監督 山口ヒロキ



私が映画監督として活動を始めてから20年以上、ずっと心の奥にあったのは「オリジナルのSFをもっと創りたい」という思いでした。しかし実写でSFを制作するには、あまりにも多くの制約があります。未来都市や架空の世界を描くには高度な技術と莫大な予算が必要です。結果として、描きたい世界観を圧縮し、観る側の想像力に委ねる表現を選ばざるを得ませんでした。

転機は2023年後半に訪れました。それまでは生成AIに関心を持ちつつも、映画制作に使えるレベルではないと感じていました。しかし年末にかけて、AIが生成する映像のクオリティが一気に跳ね上がったのを目にして、「これは近いうちに生成AIを使用した映画が作られるかもしれない」と直感しました。とりわけAIはSFとの親和性が高く、これまで諦めてきた世界観の表現が現実味を帯びてきたのです。

## 生成AI映画『generAldoscope (ジェネレイドスコープ)』の挑戦

最初に手がけたのは、6分ほどの短編映画でした。当時の生成AIは長い動画を一気に出力することができず、しかもスローモーションのようなものばかりでした。1回に2秒ずつくらいの映像でしか生成できないので、生成作業を何回も何回も繰り返し、断片をつなぎ合わせていくという、とても時間がかかる作業でした。そうした生成AIの制約を前提に脚本を組み立て、編集を加えることによって映画としての流れをつくっていきました。

試行錯誤の末に完成した作品『IMPROVEMENT CYCLE- 好転周期 -』は、いくつかの海外の映画祭に入選・受賞することができました。そして、次の挑戦として生まれたのが、オムニバス映画『generAldoscope』(2025年8月公開)の一編である『グランマレビット』です。脚本家さんの書いた脚本に基づいて、生成AIだけで本当に映画を完成させられるのか、その実験的な試みでした。

私が担当した14分の作品『グランマレビット』では、架空の都市や、細部まで作り込まれたSF世界を画面いっぱいに詰め込みました。イメージしていた風景を、初めて映像として表現し切れたという実感があります。ただし、AIを使えばすべてが簡単にできるというわけではありません。特に『グランマレビット』を制作していた2024年秋から2025年の年始の頃は、登場人物の見た目の一貫性が保てない、思い描いたカメラアングルからの映像を生成できない、後ろ姿のショットがどうしてもうまく出力されないという事が多々ありました。あるシーンでは、どうしても生成できないカットがあり、脚本やカット割りを根本から変更せざるを得ませんでした。AIは万能ではなく、試しては失敗し、考え直し、表現を組み替える必要がありました。

とはいえ、そのプロセスはこれまでの映画づくりと本質的には何ら変わりありません。制約があるからこそ、感性を研ぎ澄ませ、シビアに表現を選択するというのも同じなのです。

『グランマレビット』の制作は分業で進めました。私が映像イメージを言語化し、演出意図を細かく指示書に落とし込み、それを元に複数のアシスタントとリモートで作業を分担しました。結果的に、アニメ制作というより、漫画家の制作工程に近い感覚になっていきました。

究極的には一人でも完結できる作業ですが、仲間がいることでスピードと密度は格段に上がります。演出部出身のスタッフは、脚本を読み解き、感情や動きを理解する力に長けており、AIによる映像生成とも相性が良いと感じました。映画づくりのあり方そのものを改めて問い直す良い経験でもありました。

## 飽和の先に真価が問われる

海外の映画祭に招待していただいた中で、AIで映画を制作する多くの監督たちと交流するようになりました。韓国では2024年の段階でAI映画祭が既にいくつも開催されていました。世界中でAI映画制作のチャレンジが熾烈化していますから、遠からず、踊り場というか、飽和状態に達するのではないかと見ています。その時、自分自身の作品をどう際立たせるのか、どのような世界観を描くのかという映画監督としての原点が問われることになると思います。

私にとって、AIで映画をつくる意義は明確です。それは「クリエイティブの解放」です。頭の中にありながら、技術や予算の制約によって表現できなかったものが、ようやく外に出してかたちにすることができるようになった。抑圧されていた創造性が少しずつ解放されていく。その実感が、次の挑戦へと私を駆り立てています。

今、過去に実写で制作したSF映画をAIでリメイクする計画を進めています。当時は諦めざるを得なかったシーン描写や世界観の表現を、AI技術を使って実現したいと考えています。また、AIと実写を組み合わせたハイブリッドな映画制作にも取り組んでいくつもりです。

AIはあくまで道具です。何を描きたいのか、どんな世界を観客と共有したいのか。その問いに向き合い続けることが、映画監督の仕事だと思っています。道具が変わっても、その本質は変わりません。私はこれからも、試し、壊し、創り続けていきます。新しい表現の可能性を信じて。 (談)

## 山口ヒロキ / Hiroki Yamaguchi

映画監督

2004年、自身初の長編映画『グシャノビンヅメ』がモントリオール・ファンタジア国際映画祭でグランドプレーカーアワード銀賞を受賞。代表作は『メサイア』シリーズ、『血まみれスケバンチェーンソー』シリーズ、『トリノコシティ』など。2024年、AI映画『IMPROVEMENT CYCLE - 好転周期 -』がプチョン国際ファンタスティック映画祭を始めとした海外のAI映画祭に多数入選・受賞。現在は実写に加え、AIを使用した映画・MV・CM・プロジェクトンマッピングなどの映像を制作。また、生成AIに関するアドバイザーとしても国内外で登壇を行っている。

インタビュー実施：2025年11月12日

## 現場の現場による 現場のための AI 導入

篠田 貴之 海外戦略センター 兼 創造テクノロジー部 部次長



### 小さな取り組みを少しずつ積み重ね・・・

日本テレビの制作現場では、2016年から少しずつAIの活用が始まり、AIを用いた映像や音声のコンテンツ解析をリアルタイムに行えるシステム「AiDi (エイディ)」は数百件の社内活用にとどまらず、国内外、業界外でも導入が広がっています。

なぜ、日本テレビにおけるAI利活用が順調に進んできたのかと聞かれることがありますが、それに対する私の答えは「現場主導」の一言に尽きます。

私は、学生の頃から映像や音楽などコンテンツをつくるのが大好きで、入社してからはテクニカルディレクターとして機材オペレーションやCG制作などを含めてコンテンツ制作に広くかかわってきました。ただ、テレビ局の現場というのは安全第一主義です。放送で事故を起こすリスクを嫌うあまり、実は新しいテクノロジーを導入することに対しては慎重になる傾向があります。新しいことを試してみることが好きな私にとって、それは葛藤でした。

そこで、小さな取り組みを少しずつ積み重ねていくことにしたのです。自分でコードを書き、現場で試して、使ってもらってということの繰り返しです。最初は半分個人的な趣味のようなものでしたが、現場のペインポイント（課題、苦勞）は身をもって知っていますから自然と受け入れられ、番組でも使ってもらえるようになっていきました。

現場の、現場による、現場のためのAI導入だったことが、その拡大サイクルを回すための最も大きな原動力だったと思います。

そして、少しずつ効果が見えるようになり、理解者が増えていくにつれ、要所要所で会社による支援が加わりました。人員を増やしてもらったり、グループのIT会社に繋げてくれたり、新しい部署を立ち上げてくれたり……。ボトムアップというほどの“闘い”があったわけではなく、トップダウン的な“押し”があったわけでもなく、悲壮な危機感に背中を押されたわけでもありませんでした。

そういう意味で当社の導入ケースは特殊なのかもしれません。しかし、「AiDi」が国内そして海外の現場から多くの引き合いをいただけているのは、やはり、現場で生まれ、現場に鍛えられたソリューションだからなのだと思います。

### 次の重要テーマは「暗黙知の取り扱い方」と「人の役割再定義」

今、個人的に気になっていることが二つあります。

一つ目は、組織における暗黙知の取り扱いです。AIの進展により、これまで「暗黙知」とされてきた

曖昧で属人的な知見を、構造化・具体化・視覚化することが可能となり、これは技術の継承や業務の効率化・高度化を促進するものであり、社会全体で加速している流れだと言えます。

しかし一方で、暗黙知を過度に形式知化し、一か所に集約していくことは、新たなリスクも伴います。集中管理は活用の面では有効ですが、取り扱いを誤れば、組織の中核となるノウハウを不用意に流出させる危険性があります。また、形式化された知見に依存し過ぎることで、コンテンツ制作の多様性や創造性に偏りを生じさせる可能性も否定できません。だからこそ、暗黙知を「どこまで可視化するのか」「どのように管理・共有するのか」という視点を持ち、その取り扱いに対する意識を高めていくことが不可欠だと考えます。

もう一つは、AI時代における人の役割を再定義する必要があるということです。言い換えれば、人間とAIの責任分界点をどこに引くか。「AiDi」の主な狙いは、番組制作のベースを支えるAD（アシスタント・ディレクター）の仕事のうち時間と労力がかかる作業をAIに任せることによって、番組の質や制作の量を向上させることです。それは、決してADをAIで代替することを意味していません。ましてや、ディレクター、チーフディレクター、エグゼクティブディレクターといった役割をAIに担わせることも想定していません。番組の方針を決める中核部分までAIに渡してしまうと、社会の健全性が損なわれかねないと私は懸念しています。

人間から番組制作の能力を奪うだけでなく、もし悪意が埋め込まれたAIモデルが作られ、それに依存するようなことになると、操作された価値観や利害関係を反映したニュースやコンテンツがまき散らされることになりかねないからです。どこまでがAIで、どこからは人が担うという線引きについて、もう考え始める時が来ているのではないかと感じています。（談）

## 篠田 貴之 / Takayuki Shinoda

海外戦略センター 兼 創造テクノロジー部 部次長

2008年、日本テレビ放送網に入社。入社以来、制作技術やCGなど番組制作部門に所属し、現場の視点を生かしたAI・画像処理技術の起案・開発に取り組む。これらの成果により、「日本民間放送連盟賞」6回、「経済産業大臣賞」3回をはじめ、35件以上の受賞歴を有する。2024年より海外戦略部門で事業リーダーを兼務。

インタビュー実施：2025年12月22日

## 空気のように、感じるままに —15歳 AI ネイティブの視界

sola 15歳のAIプロデューサー



大人たちは、AIを前にして忙しい。「使わなければ置いていかれる」「仕事が奪われる」「とにかく急げ」と、どこか焦りと不安を滲ませながら、AIを“事件”として扱う。

その隣で、15歳の少年は、驚くほど静かで穏やかだ。

solaくん。2010年生まれ、徳島県海陽町と東京の2拠点で生活する中学生。生成AIを使った映像や音楽作品を次々と発表し、企業案件や万博関連の制作にも関わる。プロフィールだけ見れば「凄い少年」だが、実際に話してみると、その印象は少し違う。

AIについて熱弁するわけでもない。未来を予測するわけでもない。

「面白そうだったから触ってみた」「表現の幅が広がるから使っている」

語られる言葉は、拍子抜けするほど淡々としている。

彼にとってAIは、脅威でも革命でもない。生活の中にいつの間にか置かれていた、ひとつの道具にすぎない。

作品制作の話聞いても同じだ。

「耳鼻咽喉科のうた」——。おとなの完成では、にわかに意味が分からない（笑）。だが、なぜか引き込まれるものがある。本人はいたって平然。そこに「AIだからできた」という気負いはない。やりたかったからやった。思いついたから形にした。ただそれだけだ。

制作時間が短いことも、あっさりしたもの。



「一人で、わりと簡単にできちゃう」

大人が「効率化」「生産性」と言葉を重ねるところを、彼はさらっと通り過ぎる。AIは目的ではなく、衝動に追いつくための近道にすぎないのだ。

「ワーク」「スタディ」「ライフ」をどう分けているかという問いには、「区切っていない。考えていない。ただ、やりたいことをやっている」。AIネイティブというラベリングは大人の勝手。solaくん自身は、自分を何者かに定義しようとしなない。AIが来ても生活は変わらなかった、と言い切るその感覚に、大人はじれったさを感じるほどだ。

未来は？

「十年後？ わからないですね」「好きなことやって、遊んでほしい」

社会がどう変わるか、仕事はどうなるか。そうした問いに、彼は無理に答えようとしなない。考えないことを、怠惰とも楽観とも捉えない。まだ確定させる必要がないだけだ。

大人はAIを「どう使うか」を問う。solaくんは、「どう生きているか」の延長線上にAIを置いている。

だから騒がない。

だから急がない。

だから、自然体。

AI時代の感性は、必ずしも最先端の技術論の中にあるわけではない。

むしろ、AIを特別扱いしないこの少年の距離感の中にこそ、未来の輪郭が見えてくる。

大人が「AIだ、たいへんだ」と声を上げている間に、彼らはもう、次の遊びを始めているのかもしれない。

## sola / ソラ

15歳のAIプロデューサー

15歳中学生 / 株式会社 MUSOOO 代表取締役 / iU 客員教授 / 徳島県海陽町の中学生 / 変態ガジェット代表 / 電  
電チンドン代表 / e 場所代表

インタビュー実施：2025年11月14日

## 私の視点

# AI ×安全保障において何が重要か

松原 冬樹 Arrowheads 代表取締役



私は Arrowheads 株式会社を創業し、現在は AI セキュリティ・サイバーディフェンスソリューションを開発・提供させていただいております。AI が凄まじい勢いで進化を続ける中で、小回りの利くベンチャーは変化に対応がしやすいと感じています。そうした立場から見て、今、「AI ×安全保障領域で何が重要だと考えるか?」と問われたなら、「縦割り組織」を一度壊し、組織内の情報（コンテキスト）の流れを良くすることが重要であることを述べたいです。

AI の進化は、すでに安全保障の在り方そのものに大きな影響を及ぼしています。情報収集、分析、意思決定はもちろん、サイバー戦からオンライン HUMINT (human intelligence、人間を媒介とした諜報) に至るまで、従来は人間や組織が時間をかけて行ってきたプロセスが、急速に AI によって置き換えられつつあります。そして、ある特定の分野に限定されることなく、多発的に変化が進行しています。

人類が考案してきた概念の壁など構わずに押し寄せてくる想定外の大波に、縦割り組織が十分に対処できるようなには思えません。情報がサイロ化して横に伝わらない、人と人々が壁を越えて交わらず、コンテキストの共有ができていない。その結果、AI に関する情報や問題意識が、組織や分野を越えて共有されにくく、現場で見えている変化やリスクが、全体戦略の議論に結びつくまでに長い時間がかかり、対応がどうしても後手に回ってしまう。

先手を打つべき安全保障の分野で、結果として「かなり後方から追いかける対応」になってしまうのではないのでしょうか。

例えば、国産 LLM (大規模言語モデル) について、安全保障の観点から必要だという指摘がすでに様々なところでなされています。基盤モデルは米国大手プラットフォームに大きく先行されており、今すぐ同等レベルの性能を有する国産モデルを構築することは難しいと言われていています。ただ、安全保障の観点からは完全な他国依存は非常に危険であると考えています。もし、それが何らかの要因で使用不能になったり、シャットダウンされたりするようなことがあれば、国家的な危機を招くことになりかねません。

AI は、分野横断的にコンテキストをつなぐことにより、本来複雑である意思決定をも担えるようになります。それをを使う側の組織が分断されたままであれば、AI の強みを十分に生かすことは難しいと考えます。安全保障という国家レベルはもちろんですが、日本経済を支える個々の企業レベルにおいても「AI による大変化に対応するための組織の構造」について議論を活発化していくべきではないのでしょうか。 (談)

インタビュー実施：2025年11月28日

## 私の視点

# 研究という営みが根本的に変わりつつある

高木 志郎 Unktok 代表

私は大学卒業後に株式会社 Unktok を立ち上げ、機械学習研究を自動化する AI システムの開発に取り組んでいます。今強く感じているのは、広い意味での研究という営みの前提が、根本的に変わり始めているという事実です。

生成 AI の登場によって、研究の各工程は様変わりしました。論文はいちから精読するものではなく、要点は AI に整理してもらってつかむ。コードは人間が書くのではなく、AI に指示を与えて形にする。研究アイデアの整理や関連する膨大な研究の探索も AI が担う領域になりつつあります。研究プロセスの多くの部分が、急速に AI に委ねられ始めているのです。

仮説を立て、検証し、次の問いを生み出す——これまで人間が担ってきた工程の多くが、部分的に、あるいは連続的に自動化され始めています。「研究が速くなった」というプラスの側面の裏で、研究とは何か、研究者とは何をする存在なのかという定義そのものが揺らぎ始めているように感じます。

研究を支える制度や評価の仕組みは、依然として「人間が手作業で研究を行う世界」を前提に設計されています。論文という形式、査読というプロセス、研究費の配分、成果評価、等々、現状は AI が研究の一部を担うことを織り込んだものにはなっていません。

歪みも見え始めています。AI によってアウトプットの量が増える一方で、それを人間がすべて査読・評価することが難しくなりつつあります。これは誰かの努力不足や制度設計の失敗というより、前提としていた研究モデルと、現実とのズレが大きくなっていることの表れだと思います。

では人間はこれからの研究において何を引き受ければよいのか——。AI が問いの設計や仮説生成・検証を担うようになるなら、人間の役割は問いそのものへの意味づけ、社会的な文脈への接続といった部分に重心が移っていくのではないのでしょうか。細かい作業は AI に任せ、人はどこに向かっていきたいのかという大きな目標を決めることに集中していく。もちろんその責任を引き受けるわけですから、簡単なことではありません。その過渡期に、今、私たちは立っているように感じます。

この変化は、研究者個人の創意工夫だけで吸収できるものではありません。研究のやり方が変わる以上、それを支える制度や公的支援のあり方も、少しずつ調整されていく必要があります。

そして、何より肝に銘じなければならないことは、日本人研究者だけが内に凝り固まっていたは到底世界に太刀打ちできないということです。海外の第一線級の大学院生や研究者をもっと広く深く受け入れて、日本人研究者もそこに交わり合い、AI 時代の知の創造を日本からリードするような環境を整備すること。非常に難しい課題ですが、日本にとってはそれが第一歩だと思います。 (談)

インタビュー実施：2025年11月28日

## 私の視点

# 誰もが何でも作れる時代の創造性とは？

花光 宣尚 慶應義塾大学 特任助教



生成 AI が急速に普及し、映像、音楽、文章、ゲームといったクリエイティブの現場に広く入り込んできました。「誰でも作れる時代になった」と歓迎する声の一方で、「プロの仕事がなくなる」「AI はクリエイターの墓場」といった悲痛な叫びも聞かれます。

クリエイティブは大変な時代に入ったことは確かです。生成 AI の進化につれ、コンテンツの制作環境は激変しました。かつては専門的な技術や長い経験が必要だった工程が、プロンプト一つで成立してしまう。映像、音楽、文章・・・数時間、数日かかっていた作業が、文字通りあっという間にカタチになってしまう。AI が大量の「それらしいもの」を生み出す中で、人間が手を動かす必要性は減り、クリエイティブの現場は様変わりしつつあります。

では、「誰でも作れる」時代にクリエイターは何を創るべきなのか——。この問いに対する明解な答えはまだありません。しばらくの間行ったり来たりの模索が続くと思います。ただ、一つ言えるのは、実際に触ってみた肌感覚、自分の身体を使った本物の体験、人と人との交わりの中で心から湧き出る喜怒哀楽の情、知識の枠をはみ出た所で起こる セレンディピティによる気づき——そういった（今のところ）人間にしかできないプロセスが、これまで以上に大切になってくるということです。

そもそも、人間の創造性とは何なのでしょう。例えば、芸術作品を鑑賞している時、ぼんやり音楽を聴いている時、歩きながら仕事のことを考えている時、突然、頭の中で何かがスパークして今まで考えもしなかったものがふっと浮かび上がってくる、そんな経験をしたことはありませんか。五感からのインプットが交じり合った時に一種の“化学反応”が起き、自分の中に何か新しいものが生まれてくる。まるで「夢」でも見ているかのような・・・。

自分が経験したことの組み合わせからできたものですから、自分にとってとても大切に、糧となるものであるようなありがたさ。誰かに指示されたものでも、コンピューターが計算したものでもない、確かに「自分」の内から生み出されたもの。言葉ではうまく表せないのですが、クリエイターと呼ばれる人たちに限らず、誰もが味わったことのある心の記憶なのではないでしょうか。

AI 時代の入り口に立った今、新しいものを創り出すという人間ならではの行為をもう一度見つめ直し、その意味を問い直すことが必要になっています。人間のクリエイティビティはまだほんの一部しか開花していない。AI は人間から創造性を奪うものではなく、人間の潜在能力を引き出し、創造力をさらに拡張するものであってほしいし、そうしなければならない。そのためにはどうすればいいのか——それが私にとって最大の研究テーマです。

(談)

インタビュー実施：2025年11月25日

AI ガバナンスを巡る論点 2025

鼎談 デジタル政策フォーラムのアクション

# 多極化する世界における AI ガバナンスを追求

AI を巡る状況は、経営、経済、社会、文化、政治、外交、安全保障など様々な局面で急速に変化している。デジタル政策フォーラム（DPFJ）の認識と対応はいかに——。谷脇康彦、中村伊知哉、菊池尚人の三名が語り合った。

## Discussants

### 谷脇 康彦

デジタル政策フォーラム 代表幹事  
株式会社インターネットイニシアティブ 代表取締役社長執行役員

### 中村 伊知哉

一般財団法人デジタル政策財団 理事長  
iU 学長

### 菊池 尚人

デジタル政策フォーラム 代表幹事代理  
慶應義塾大学大学院 メディアデザイン研究科 特任教授

※デジタル政策フォーラムの Web サイトにも掲載している

<https://www.digitalpolicyforum.jp/>

# AIをめぐる現状認識



**谷脇** デジタル政策フォーラムとして『AI ガバナンスの枠組みの構築に向けて (Ver 1.0)』を公表したのは2024年7月。まずは論点を整理することから始めた。同年12月に『AI ガバナンスの枠組みの構築に向けて (Ver 2.0)』を公表。日本政府がAI法の制定に動き出していたところに、法律を作るならば、「基本法」とし、厳しい規制をかけるべきではないと提言した。結果、AI法はほぼ提言通りの中身となった。これは評価できる内容だ。

今回 Ver 3.0 をまとめていて強く感じたのは、**法的な枠組みのあり方論から具体的な政策論に焦点が移ってきたこと、そして政策という範疇を超え、文化、宗教、民主主義**

**といった領域にまで論点が広がってきていることだ。**

これは、多くの人がAIを実際に使い始めたからだろう。AIを実際に使ってみて、「AIのあり方」ということについて具体的に議論できる環境になった。実利用が進むほど論点が広がり、鮮明になっていく。

例えば、リスクへの備え。米国ではトランプ政権がAIを極力規制しない方針を示している。自国のAIプラットフォームが競争優位にあるからだ。これに対して、欧州や韓国ではリスクの強弱でAIに対する規制の度合いを変える方針だ。AIを安心して利用するために規制が不可欠という考え方からだ。個人的には「AIのリスク」を客観的・定量的に評価・定義するのは非常に難しいと思うが、AIリスクへの対応が国によって大きく違ってきている。米国内でも連邦政府はAI放任主義だが、ほとんどの州でAI規制法が成立しており、軋轢が生じている。ハードローでいくのか、ソフトローでいくのか、考え方に大きな違いがある。

強いAIを生み出しているのは米国のプラットフォームであり、AI前に構築されてきたプラットフォームの市場独占力、データ独占力というものが、AI後にもそのまま踏襲されている。競争政策と産業政策の観点から丁寧に議論すべきテーマである。

**中村** 提言・ステートメントというものは三つも作ると大体同じような内容に収束して役目を終えていくものだが、今回を含めた三つのAIステートメントは、それぞれに重要な意味があった。

Ver 1.0では、「AIのコントローラビリティ」について問題提起した。Ver 2.0では「AI法のあり方」について提言した。そうしたら実際そういう法律ができた。Ver 3.0ではテーマごとに深掘りして具体化すればいいのかなと思っていたら、むしろ論点が広がった。

要因の一つは想定以上のスピードでAIが進化し、影響を与える領域が拡大していること、もう一つはトランプ政権が世界の場面（ステージ）を変えてしまったということ。関税にしろ、ベネズエラへの軍事侵攻にしろ、グリーンランド買収の件にしろ、過去一世紀以上をかけて形作ってきた自由資本主義の定石が通用しなくなっている。自由民主主義の旗手だった米国が専制主義側に回ってしまったようにも見えるし、法の支配という近代国家の拠り所さえ揺らいでいる。そういう不安定な政情の上に、**強大なパワーを秘めたAIを誰がどのような仕組みでコントロールするのかという論点**が乗ってきた。

さらには、ハードロー派の EU と放任派の米国が相反する構図の中で、ソフトロー主義を取っている日本のポジショニングが適切なのか、今後もそれで良いのかという論点も新たに浮上している。

今回の「AI ガバナンスに関する提言 Ver 3.0」はそうした混沌とした状況に向き合った結果である。

**菊池** 「AI × 政治」「AI × 経済」については、今回の「AI ガバナンスに関する提言 Ver 3.0」をまとめるに当たっても、第一線で活躍する様々な領域の研究者に論じていただいた。

ここからは、教育であるとか宗教であるとか、次世代にどう紡いでいくか、人の心のあり様はどうあるべきか、といったテーマをしっかりと考えていく必要がある。

そのような議論では、世界における日本のポジションを直視することから始めなければならない。30 年前の日本は世界第 2 位の経済大国だったが、2025 年にはインドに抜かれて GDP（国内総生産）で第 5 位に後退する見込みで、今後さらに順位を落としていく。経済的な中位国として、どのような戦略で AI をとらえるのかという意識を持たなければならない。

**谷脇** 議論の前提として「AI = LLM（大規模言語モデル）」と考える人が多いが、ルーターが相互に自律的につながり合った分散型ネットワークであるインターネットに照らしてみると、LLM のような大規模・集中型 AI だけではなく、小規模・分散型 AI の可能性も考慮しておくべきだと思う。分散した AI が仮想的に、あたかも一つの AI のように動くイメージだ。その一つのカタチが「パーソナル AI」だ。

日本の勝ち筋は「パーソナル AI」のような分散型 AI ではないかと思う。集中型と分散型が共存して互いに競い合うことによって、AI の世界に健全な競争が生まれ、ひいては健全なマーケットに成長するのだと思う。

**中村** AI を取り巻く環境だけでなく、AI そのものも急速に変化している。「LLM = AI」という考え方で論じられてきたが、そろそろ世界中のオープンになっているデータは食い尽くすということになると、個別の組織が持っている良質なデータに基づく分散型 AI やパーソナル AI が勃興してくる。その変わり目にこそ日本の勝ち筋も見えてくるかもしれない。2026 年には、AI が身体性を持つ、つまりロボットと合体してフィジカルに現実空間に出ていくということが本格化していこう。そこでも、競争の局面が大きく変わる。

我々は目の前にある AI を前提にして、経済、社会、文明にどのような影響を及ぼすかっていうことを掘り下げてきたわけが、AI は凄まじいスピードで進化している。シンギュラリティの到来もそれほど遠くないかもしれない。AI そのもののが劇的に変わり続けるスピードに、法体系だとかコントローラビリティだとかの議論が追い付けるのか、AI をいかに論じるかという論点も浮かび上がってくる。

## 有識者・企業・クリエイターへの ヒアリングを踏まえた気づき

**谷脇** 企業の経営層、リーダー層の AI に対する理解が非常に進んだというのが、Ver 2.0 からの最大の変化だと感じる。「AI を使うか使わないか」から「AI をどう使うか」に経営戦略の焦点が明らかに移った。日本の DX（デジタルトランスフォーメーション）は経費削減を目的とすることが多かったが、AI 導入に関しては最初から新規事業創出、付加価値向上を目指すケースが多く、その違いが鮮明であることが印象的。

特集「AI × クリエイター」の事例からは、これまでに確立されてきた領域・分野が AI によって根底か

ら変わりつつあること、様々な挑戦者が登場していること、その一方で既存プレイヤーの中には事業継続が困難になりつつある人たちもいること、等々、悲喜こもごもの様相を垣間見ることができた。人間のクリエイティビティの根源的なところが問い直されているという指摘も非常に重い意味がある。クリエイティブ以外の領域も含めて横断的な調査と分析を重ねることによって、さらに厚みのある議論になるはずだ。



**中村** AIは経済・社会の基盤になりつつある。企業では「効率化」よりも「戦略の基盤」として認識されている。そして、経営層のリーダーシップが顕著だ。AI利用度の差は、ひとえに経営者のリテラシー差の表れということになってきた。

一方、文化面の議論も盛んになってきた。コンテンツ、メディア、ジャーナリズムについての問題提起、クリエイティブの現場の声・・・AIの議論がいろいろな分野に広がっていることを実感した。

そして、メタ認識の議論が盛んになってきた。宍戸常寿 東京大学教授は民主主義について問いかけ、國領二郎 慶

應義塾大学名誉教授は社会構造の変貌を説き、須藤修 中央大学教授は文明論を語った。非常に大きな問題提起が表に出てきている。今の国際的な変動とAIの衝撃が相まって、根源的なところを問い直すことが求められているのだと思う。そして、大澤淳 中曽根康弘世界平和研究所上席研究員の「経済合理性よりも安全保障を優先すべき」という提言、村上明子 AISI 所長の「AIガバナンスとは安心を支えるプロセスである」という指摘は、いずれも非常に重いものがある。

「AI×クリエイター」については私も深くかかわっているが、悩ましい問題が浮上している。AIを使わないとクリエイティブが成り立たなくなるという認識が広がる一方で、ユーザーやファンの中にAIでコンテンツを作ることに對する嫌悪感が生まれ、AIヘイト論が世界的に広がっている。僕はAI推進派なのでそういうことをSNSに投稿したりすると大炎上してしまう。コンテンツ企業のなかには「コンテンツ制作にAIを使わない」と宣言するところまで追い込まれたところもある。一過性の現象なのか、コンテンツ制作へのAI活用を躊躇することが常態化するのか・・・。先が見通せない。

**菊池** 企業がAI利活用を進めていくと必ずぶつかるのが「データ」の壁である。自社が蓄積してきた様々なデータをAIに学習させることが独自の付加価値となり競争力を生み出すのだが、肝心のデータが紙ベースのままだったり、フォーマットがばらばらで構造化されていなかったり、有効に生かされていない状況がまだまだ見られる。今回のヒアリングでも複数の企業、団体からそうした問題意識が聞かれた。

これは個社だけではなく、中小企業同士の連携や公的データの利活用、産業間連携、国際連携といった重層的な課題でもある。国立循環器病研究センターの西村邦宏先生は、日本の医療データの収集・整備・活用がなかなか進まないことに強い危機感を隠さなかった。日本的な安全・安心志向がデータの利活用においてもチャレンジを許さないという指摘は非常に重い。デジタル庁は「誰一人取り残されない、人にやさしいデジタル化を。」というミッションを掲げているが、果たしてこの理念のままで良いのだろうか。明治の文明開化は日本が生き残る手段であり、目的でもあった。

社会構造をそのままにDXやAIの利活用を推進しようという考え方には限界がある。人口が減少し

てAIを使わないと成り立たない社会になるので、それに合わせて社会構造を思い切って変える決意で取り組まないと進まない。世界から大きく遅れている日本の中で議論しても取り残されるだけだ。海外の先頭集団は今どこにいるのか、どこに向かっているのかをしっかりと認識すべきだと感じる。

**谷脇** サイバー空間にどこまで国際法が適用できるのかという議論がある。日米欧はリアル空間と同じように適用できるという考えである一方、中国・ロシアはそれとは違う意見で、なかなかまとまらないという状況があった。ところが、最近の米国の行動を見ると、そもそも国際法そのものが機能するのかという話になってきている。

世界経済フォーラム（WEF）の「The Global Risks Report 2026」の中で最も目を引いた表現は、「Multipolarity without multilateralism（多国間主義なき多極化）」<sup>1</sup>だ。**多国間主義、すなわちルールに基づく国際秩序が衰退し、国境を越えた経済的・軍事的紛争のリスクが高まっている**という分析で、まさにその通りだと思う。デジタル技術がその駆動力になっているのは明らかで、AIの普及がそれに拍車をかける。アルゴリズムは妥協ということを知らないので、多極化がさらに加速していく恐れがある。

かなり深刻な状況に来ているのだが、危機的な状況を回避するためにどうすればいいのかについては、個別領域ごとではなく、**マルチステークホルダーで議論すべきである**。しかし、マルチステークホルダーをつなぎとめる吸引力が全く見当たらない。皆が同じ方向に向かっている場合、マルチステークホルダー主義で議論するとまとまりやすいが、全く逆の方向を向いている時、つまり多極分散状態の場合にはマルチステークホルダーによって結論を導き出すことは困難だ。

**中村** 日本は著作権法を改正してAIによる学習をほぼフリーにした。AIの開発を促し、コンテンツ制作を強化していくという産業政策的な狙いがあった。しかし、そのようなことを公に言うと必ず炎上する。海外から批判のつぶてが飛んできたりもする。すべての案件がグローバルで語られるので、国ごとに閉じた議論がそもそも成立しなくなっている。

いろいろな人からそれぞれの意見を聞くのも大切だが、ここまで多極化してくると、巨大プラットフォームの戦略や考え方の真髄に迫ることが本丸のように思えてくる。米国はもちろんだが、その対極にいる中国の研究者や企業のフロントランナーがこの局面をどう見ているのかも気になるところだ。

**菊池** トランプ米大統領の手によって、まるで19世紀後半の帝国主義に巻き戻った感さえある。**国際秩序という共同幻想が消えた混沌の中で、米国、中国という巨大プレイヤーやグローバルサウスがどう動くのか、注視することも必要だ**。

## 「AI ガバナンスに関する提言 Ver 3.0」 の意義と今後

**谷脇** 「AI ガバナンスに関する提言 Ver 3.0」の構成は「Ver 2.0」からほとんど変わらず、書き足しに終了した。その意味で、今まで積み重ねてきた議論の方向性は間違っていなかったことを確信した。産業政策、競争政策、安全保障、コンテンツ、文化、宗教・・・と議論すべき範囲はとてつもなく広がった。しかも、それぞれが相互に深い関係性を持っている。

1 <https://www.weforum.org/publications/global-risks-report-2026/>

ただし、政策的には後手に回っていると云わざるを得ない。AI法（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）が2025年9月に施行され、それを受ける形で12月には「人工知能基本計画」が閣議決定された。残念ながら、各省庁の政策の寄せ集めに終わっている。もっと俯瞰的で相互関連性を重視した政策を組んでいかないと、日本のAI活用が進まない。

**省庁間の壁、業態間の壁を破り、高い次元での政策論に持ち込むためにはどうすれば良いのか**という、根本的な課題に本気でぶつかっていかねばならない。

**中村** 国際情勢が様変わりして、米国と欧州の狭間にいる日本の立ち位置が明確になった。同時に民主主義とか資本主義の基盤が大きく揺らぎ始めた。AIの軍事利用という重大テーマも浮き彫りになってきた。認知戦やサイバー攻撃、さらには兵器利用にまでAIは深く入り込んでいる。近代国家の存立を危うくする恐れさえ出てきた。

抽象論だけではだめ、かといって細かい個別領域ごとでの議論だけでもだめ。**鳥の目をもって全体を俯瞰し、要所にある壁を破らなければならないが、それは日本が非常に不得手とするところだ。だが、もはや、その問いから逃げていられる状況ではなくなっている。**

規制と振興の両面から思い切った打ち手を繰り出していかねばならない。そして、**特に教育、医療、行政におけるAI利活用を加速することは、最重要政策課題**として取り組むべきだ。



**菊池** 各省庁縦割りであることに加え、施策が短期視点であることが問題だ。メタ視点の議論が必要であるということは今回の「AIガバナンスに関する提言 Ver 3.0」の非常に重要な投げかけになっていると思う。

**谷脇** 今回、デジタルデモクラシーについても触れた。AIが普及していくと民主主義はどのように形を変えていくのか、民主主義は機能するのか——。極めて重要な論点である。

もう一つ、重要な論点としてサイバー空間における地政学がある。端的に言えば、**サイバー空間の中に国境ができて始めていることをどうとらえればいいのか**ということ。インターネットは時間と距離を超越する、サイバー空間には国境がないと言われてきた。現実世界が国境で区分され多極化したとしても、サイバー空間では繋がりが合っているということが一つの救いだったが、今、そのサイバー空間までもが分断されようとしている。世界の分断がいよいよ決定づけられてしまうのではないかと深く危惧している。

デジタル政策フォーラムとしては、領域横断で定点観測を続けるとともに、様々な領域の人たちとの対話を継続することによって、俯瞰的視点からの議論を触発する役割を果たしていきたいと思う。

**中村** AIが国家のあり方や安全保障という領域に入り込んでくるのなら、政治のど真ん中にある政治家がこれをどうとらえているのか聞いてみたいし、そうした議論を始めなければならない。

人々がAIに何を求めるのかという点も面白いテーマだ。電通が2025年6月に実施したアンケート「対話型AIとの関係性に関する意識調査」<sup>2</sup>によれば、「対話型AI」に感情を共有できる人は64.9%で、「親友」

2 <https://www.dentsu.co.jp/news/release/2025/0703-010908.html>

(64.6%)、「母」(62.7%)に並ぶ”第3の仲間”になっているという。特に若い世代ほどAIに対する愛着や信頼が高い。そのうち、「AIと結婚する」という人も増えてきて、日本の少子化をさらに加速させてしまうかもしれない。

これまで目指してきたAIと、今後みんなが欲しがるAI像は違ってくるかもしれない。AIは先生なのか、上司・同僚・部下なのか、友達なのか、恋人なのか、はたまた神なのか——。国民性によっても大きく違ってくるだろう。興味深いテーマであり、それによって我々が立てるべき問いも変わってくるだろう。

**菊池** AIステートメント Ver 4.0では、**宗教、結婚、家族、コミュニケーション**といったテーマに迫ってみたい。人間はAIによって救われるのか、AIによってもっと理解し合えるのか、AIによって愛のカタチ、社会のカタチはどう変わるのか——。とても興味がある。

**谷脇** 近代国家を構成する立法、司法、行政の三権の中で、例えば司法の機能をAIに置き換えたらどうなるのかをシミュレーションしてみるのも次の課題を見出していくうえで有用かもしれない。AIを起点にこの世の中の機能がどこまでAIによって代替できるのかを考えることで、逆にAIの持つ限界、人間との共生の姿が見えてくるかもしれない。引き続き広範な議論を巻き起こしていきたい。

**中村** 経営戦略から国家論、文化、哲学まで様々なことを、みんなで自由に議論できる時代に生きられて本当に幸せに思う。1000年に一度めぐってくるかどうかの稀有なチャンスだ。AIよ、どうもありがとうと言いたい。AIステートメント Ver 4.0が今から楽しみだ。

鼎談実施：2026年2月2日

# おわりに

---

デジタル政策フォーラム（DPFJ）が2024年春にAIガバナンスの議論を開始してから、季節が巡るごとにその風景は劇的に塗り替えられてきました。2024年7月の「AIガバナンスの枠組みの構築に向けてVer 1.0」、同年12月の「同 Ver 2.0」を経て、今回お届けする「AIガバナンスに関する提言 Ver 3.0」は、単なる情報のアップデートを越え、AIが私たちの文明そのものを根底から揺さぶり始めたことを告げるマニフェストとなりました。

通常、3度目の提言ともなると、総論から各論に移り、議論は収束に向かうものですが、逆に論点が広がり、さらなる総論へ拡張した感があります。

Ver 1.0ではAIの「コントローラビリティ」を問い、Ver 2.0では「AI法のあり方」について規制を最小限に留めるべきだと提言しました。その後、日本はAI法を策定し、実装に至ります。Ver 3.0ではテーマごとに論点を掘り下げることが見込まれました。

しかし、今回のとりまとめで見えてきたのは、法的な枠組み論が「序章」に過ぎなかったという現実です。議論は今や、文化、宗教、民主主義、さらには人間のクリエイティビティの根源といった、文明論的な領域へと拡散し、深化しています。

要因は二つあります。一つはトランプ政権が世界のステージを変えたこと、もう一つは想定以上の速度でAIが進化し、領域が拡大していることです。

米国による関税にしろ、ベネズエラへの軍事侵攻にしろ、グリーンランド買収の件にしろ、過去1世紀以上をかけて形作ってきた自由資本主義や民主主義の定石が通用しなくなっています。法の支配という近代国家の拠り所さえ揺らいでいる情勢です。

こうした不安定な政情の上に、さらに強大な力を秘めるAIを誰がどのような仕組みでコントロールするかという問いが乗ってきました。ハードロー派のEUと放任派の米国が相反する構図の中で、ソフトロー主義を取る日本のポジショニングが適切なのかという論点も浮上しています。

大規模・集中型のLLM=AIという考え方で論じられてきましたが、そろそろ世界中のオープンなデータは食い尽くされ、個別の組織が持つ良質なデータに基づく分散型AIやパーソナルAIが勃興してくると見られます。そこに日本の勝ち筋も見えてくるかもしれません。

AIが身体性を持つ、つまりロボットと合体して現実空間に出ていくということも本格化していくでしょう。そこでも競争の局面が大きく変わります。AIそのもののが劇的に変わり続けるスピードに、法体系やコントローラビリティなどの議論が追い付けるでしょうか。

企業へのヒアリングを通じて、AIが「効率化」のツールから「経営戦略の基盤」へと昇華したことを実感しました。しかし企業の利用格差も顕著です。AI利用度の差はひとえに経営者のリテラシー差の現れとなっています。

また、文化・クリエイティブの現場では、より複雑な感情が渦巻いています。生成AIが表現を解放する一方で、クリエイターの苦悩や、一部ユーザの生成AIへの拒否感のはざまに、生成AIがどこに腰を落ち着けるのかが見通せません。

有識者からは、メタ認識の議論が提起されました。民主主義や文明論、安全保障が重いテーマとなっています。

AIの進化は、私たちに「人間とは何か」という根源的な問いを突きつけます。最新の意識調査によれば、若い世代ほどAIを「親友」や「親」のように信頼し、感情を共有できる仲間として受け入れ始めています。AIが結婚相手やパートナーとなれば、少子化や家族のあり方といった社会の基本単位も再定義されるでしょう。

AIを起点にして、近代国家のシステムや個人・社会のあり方を含め、世の中の事象をメタ視点からアップデートする機会が訪れています。抽象論だけでは不足し、とはいえ細かい個別領域ごとの議論でも全体像が描けない。

「鳥の目」を持って全体を俯瞰し、要所にある壁を破らなければならない。省庁の壁や業態の壁。経済と社会の溝。作り手と使い手のはざま。それらを溶かさなければならない。けれど、それは日本が不得手とするところです。どう向き合うのがいいでしょうか。

人類の能力を超えるAIという技術。このように、今、AIを媒介にして、経営戦略から国家論、社会、技術、文化、哲学まで、みんなで自由に議論がなされています。あらゆることが一斉にバージョンアップされる。これは産業革命を超えるインパクトを持ちます。1000年に一度めぐってくるかどうかのチャンスです。この時代に生きられることを、私は本当に幸せに思います。

AIよ、どうもありがとう。

本提言の作成にあたり、多忙な中インタビューや議論にご協力いただいた、有識者、企業関係者、クリエイターの皆様に、心より感謝申し上げます。

Ver 4.0 が今から楽しみです。

中村 伊知哉  
一般財団法人デジタル政策財団 理事長  
2026年3月

## AI ガバナンスに関する提言 Ver 3.0

発行者 デジタル政策フォーラム  
編集・制作 合同会社デロイトトーマツ  
特別協力 一般社団法人 CiP 協議会  
発行 一般財団法人 デジタル政策財団  
発行日 2026 年 3 月 16 日