NEWSLETTER

NIALS における「AI ラウンドテーブル」の概要報告

－AI をめぐる制度設計への示唆－

ロボット/AI & 独禁/通商・経済安全保障ニューズレター　　2024 年 9 月 30 日号

執筆者:
藤井 康次郎
k.fujii@nishimura.com
角田 龍哉
t.tsunoda@nishimura.com

EU では、2024 年 7 月、世界初の AI に対する包括的な法規制である AI Act が公布され、日本でも内閣府の AI 制度研究会が今秋中の中間とりまとめを予定しています。西村高等法務研究所（Nishimura Institute of Advanced Legal Studies、NIALS）では最先端の理論と実務の架橋に向けた様々な研究活動を行っていますが、こうした状況を見据え、2024 年 5 月、OECD や UNESCO 等で AI 政策に従事されていた Ms. Sasha Rubel[1]（現在は Amazon Web Services のヨーロッパ・中東・アフリカ地域における AI 公共政策責任者）を講師にお招きし、弊所弁護士によるモデレーターの下、経済、競争、知財、行政等の分野の第一線でご活躍されている日本の研究者の方々とともに「AI ラウンドテーブル」を開催しました。

本稿ではその概要をご紹介いたします[2]。

1. AI 規制に関する講演概要（Ms. Sasha Rubel によるプレゼンテーション）

- AI に関する規制の対象は、必要十分かつ透明な形で規定される必要がある。例えば、AI Act 提案当初の「AI システム」の定義では、スプレッドシートやウェブ会議システム等のあらゆるソフトウェアが含まれ得る内容になっていたため、新しい規制がかえって AI に対する不透明性・不確実性を増幅させるおそれが議論されていた。また、すでに他の EU のデジタル規制が小さいがスキルのある事業者のコンプライアンスコストを過剰に高めている面があるが、AI のイノベーションに配慮した規制内容とすることで、そうした状況を避ける必要があることも議論されていた。
- AI にリスクがつきまとうことは否定できない一方で、AI Act が導入されたからといって AI のリスクがゼロになるわけではない。AI に対する規制の設計においては、AI のリスクは、通常、その設計時点で判明しているユースケースと結び付いているに過ぎないことや、特に生成 AI の利用については未だ黎明期にあることに留意する必要がある。AI Act が完全施行される 2 年後の技術水準やユースケースを現時点で見通すことには、（AI Act の議論が始まった 2021 年時点で現在の生成 AI の隆盛を一般には予見し難かったことと同様に、）引き続き困難があるというある種の限界が、その規制の設計には織り込まれている必要がある。同様に、発展途上国における AI の利用のほか、統計的ソリューションの提供、がんの早期発見、子どもの特性に応じた教育における利用等の AI のユースケースがもたらす社会的なメリットとのバランスにも配慮する必要がある。

---

1 https://oecd.ai/en/community/sasha-rubel

2 東北大学大学院法学研究科 伊永大輔教授、東京大学法学大学院政治学研究科 滝澤紗矢子教授、慶應義塾大学商学部 久保研介准教授、東京大学大学院法学政治学研究科 巽智彦准教授、東京大学大学院情報学環・学際情報学府 酒井麻千子准教授（順不同）にご参加を頂きました。ただし、本稿はご参加頂いた先生方のご意見等を示すものではなく、本稿の責任は執筆者にのみあります。

- 責任ある AI の構築に向けた検討課題の例として、AI モデルの公正さがある。例えば、男性の学習データセットのみを用いたモデルで女性用の健康診断を行わないようにするためには、データインクルーシブが必要になるが、どの程度のバランスをとる必要があるかについては、AI モデル開発・利用の文脈に即して「公正」の意義を検証する必要がある。
- AI のアウトプットの安全性（セーフティ）に関しては、引用された内容である旨が明確に表示されているが中傷的な内容には警告も表示すべきか、意見論評である旨が明確に表示されているが中傷的な内容についてはどのように考えるか、特定の個人に関する医療的、法的、政治的、金融・財務的な質問や武器兵器に関する質問は除外すべきか等の実装上の課題がよく見られている。
- リスクベースアプローチは、リスクの程度だけでなく性質に応じて異なる規制アプローチを採ることも含意した規制アプローチを意味している。例えば、おすすめの音楽を選定するために用いる AI と、レントゲン写真上で腫瘍を特定するために用いる AI とでは用いるべきアプローチやガードレールが異なるといった、具体的な用途ごとの影響に焦点を当てながら、AI に対する規制設計を検討する必要がある。
- EU における調査の一つでは、①調査対象の従業員のうち 86%が 2028 年までに大半の組織が AI を利用すると予想していること、②生成 AI が働き方を転換すること、③AI スキルの獲得が従業員の給与の増大や他のキャリアの可能性の創出につながること、④AI スキルのある労働力がもたらす生産性は大きなものになること、⑤トレーニングプログラムの認知度の向上を通じて AI スキルの格差は減少し得ることが明らかにされている [3]。
- EU における別の調査では、調査対象の EU 市民の過半数が AI は日々の業務に好影響を与え、半数が安全やセキュリティに対しても好影響を与えると回答している一方で、35%が AI によって行われている決定処理のプロセスを理解できていないことが明らかにされており、政府機関と AI 関連事業者が連携して市民による AI の責任ある利用を奨励しつつ、AI がもたらす便益とリスクに対する公衆のアウェアネスを高める対応策が指摘されている。また、欧州企業の 21%が、デジタル技術に関連するコンプライアンスや法的な不確実性を、デジタル技術がビジネスに与える影響の障壁として挙げている [4]。

2. ディスカッション概要

- AI の学習リソースとなるインフラクラウド間では、（EU データ法にかかわりなく）顧客によるサービス間の乗り換え（スイッチイング）のために必要な施策がすでに各社によって講じられている。米系や EU 系、中華系のインフラクラウド間での競争があり、規制の設計上も、そうした実態の的確な把握・認識が第一に重要となる。
- AI Act では当初 AI のリスクを 4 類型に分類していたものの、技術発展によって、汎用目的 AI システムやシステミックリスクといった概念・類型が早々に追加されるに至った。こうした修正は、必ずしも悪意のある事業者がそうした技術を開発したことによってもたらされたわけではなく、技術発展にキャッチアップするために行われたものである。そして、イノベーションや技術発展が活発な領域において、事前に技術発展のあり方を先取りしたり、制裁だけに主眼を置いたりした規制設計が困難であることを示唆しているように思われる。また、リスクの分類それ自体についても、例えば、AI を利用した GPS シ

---

[3] “ACCELERATING AI SKILLS, PREPARING THE WORKFORCE FOR JOBS OF THE FUTURE” <https://assets.aboutamazon.com/bb/2e/9077b9f44a2898c01fcc7f35440d/aws-ai-europe.pdf>.

[4] “UNLOCKING EUROPE’S AI POTENTIAL IN THE DIGITAL DECADE” <https://www.unlockingeuropesaipotential.com/_files/ugd/c4ce6f_ecf071799e4c4eba80113648d2b1090b.pdf>.

ステムを組み入れたデバイスを使って、救命救急の際のルートを特定しようとする場合には高リスク AI になるのか等、実際には個別具体的な用途や事例ごとの検討の必要があることから分かるように、分類自体で具体的な問題が解決するわけではない。

- すべての AI 関連のサービスが生成 AI によって行われることを制度設計の前提にする必要はなく、従来の予測、識別等の AI によって引き続き処理できるユースケースも十分ある。関連して、クラウド上の計算資源の消費量等についても、あらゆる AI に関する利用が高度な生成 AI モデルの構築のための学習のように大規模な計算資源を必要とするわけではないので、KYC や環境対策の一環としてそのすべての利用状況をサービス提供者が把握しておく必要があるというわけでもない。
- AI Act の域外適用の範囲（AI Act 2 条 1 項）は GDPR と同等以上に不明確な部分があるように思われ、ガイドラインの制定が待たれる。例えば、AI モデルの開発者の立場からすれば、AI 実装者（deployer）がどの法域に所在し、どの法域でサービス化するかを厳密にコントロールすることは通常できないので、AI モデルを開発する時点では当該 AI モデルを組み入れた AI システムに AI Act が適用されるかは不確実なことがある。また、例えば、日本にいる EU 市民が AI 医療機器によって診断を受けた後、EU 域内に移動したような場合に、どのような事情があればどのような点で AI Act が適用されるか等の個別具体的なユースケースごとの検討が必要になってくる。
- 第三国から見て AI Act に技術的な輸出障壁になる規定があるとすれば、AI Act の域外適用の範囲をめぐる論点は、同時に、国際協定（貿易の技術的障害に関する協定（TBT 協定）等）上の論点にもつながり得る。欧米間についていえば、貿易技術評議会（Trade and Technology Council）等といったこうした通商関係の論点についても議論できる場があり、機能している。また、いわゆる域外適用の範囲（国際的適用範囲）は、法益の保護とはやや毛色の異なる競争条件のイコールフッティングという観点からも議論されることになり、議論をさらに複雑化させる可能性がある。
- AI Act 上、AI オフィスによって提供されるテンプレートに沿って、汎用目的 AI の学習に利用されたコンテンツに関する十分に詳細な概要を策定し、公表する義務（AI Act 53 条 1 項（d））を課す主たる目的は、学習データ上のバイアスに対する透明性を確保し、これを制御するというよりは、著作権者が EU 著作権指令 4 条 3 項（及び同指令に基づく EU 加盟国法）上のテキスト・データマイニングに対するオプトアウト権を有していることを前提に、その適切な行使に資する点にあると思われる。このテンプレートは、営業秘密の保護や業界慣行に即した形で準備されることになると見込まれる。
- 生成 AI といわゆる自社優遇（自社サービス上で自社関連のサービスを優先的に取扱う措置）の懸念との関係では、生成 AI モデルの開発者としての対応と、生成 AI モデルの実装者としての対応とで分けて議論をする必要がある。生成 AI モデルの開発者の観点からは、（自社ないし自社がコントロールする事業者の商品役務に関するデータと、その他の事業者のデータの取扱いの関係を含めて）学習モデルの公平性の確保や、バイアスの排除という観点が重要になる。実際、一部のモデルでは、どのようなトレーニングデータを用いて開発したものかや、想定されるモデルのユースケースや限界等の情報を公開している取組みが見られる。また、生成 AI モデルの利用・実装の観点からは、様々な AI モデルを比較・選択できる環境の下で、ソリューションアーキテクトやパートナーによるサポート、実装者自身のデータセットを用いた（関連性・正確性等のカスタマイズを含む）モデル評価機能を活用することを通じて、（上記のような懸念が当てはまるかを検証しながら）最適なモデルを選択できることが重要になる。

日本では、AI ガバナンスポリシーの策定の奨励を含む AI 事業者ガイドラインが策定され、その実効性の確保措置や AISI による安全性評価の検討とともに、特に影響大・リスク大の AI 開発者に対しては（法制度による対応の可能性を含んだ）確実な対処が志向されています。EU AI Act の経験が示す、①規制対象の明確

性、②技術発展やメリットを動的に取込む仕組みの必要性、③著作権をはじめプライバシー、セキュリティ、競争[5]といった他法益との相互作用に基づいた必要十分な範囲での設計の必要性、④国際協定や DFFT のような他の国際的イニシアチブ、国際標準との整合性を含めた国際ルール形成のリードの重要性といった観点は、日本の AI 関連事業者が、こうした各取組みをめぐる政策動向をフォローし、そのプロセスへの参画を検討する際に参考とできる着眼点になると思われます。

西村あさひ 広報課 newsletter@nishimura.com

[5] AI Act 上、市場監視規則（（EU）2019/1020）34 条 4 項に基づく報告義務の一環として、製品の安全性等を所管する市場監視当局は、市場監視活動の過程で確認された競争に関する EU 法の適用に潜在的な関連性を持ち得る情報を、欧州委員会及び関連する各加盟国競争当局に毎年報告する義務、及び、毎年、その年に発生した禁止された慣行の利用及びそれに対して講じられた措置についても欧州委員会に報告する義務を負うと定められています（AI Act 74 条 2 項）。