第4回　AI制度研究会　議事要旨

1.日　時　　令和6年9月12日（木）16:00～18:20

2.場　所　　中央合同庁舎8号館1階 講堂

3.出席者

○ AI戦略会議　構成員

座　長	松尾　豊	東京大学大学院工学系研究科 教授
構成員	岡田　淳	森・濱田松本法律事務所 弁護士

○ AI制度研究会　構成員

座　長	松尾　豊	東京大学大学院工学系研究科 教授
座長代理	村上　明子	独立行政法人情報処理推進機構AIセーフティ・インスティテュート 所長
構成員	生貝　直人	一橋大学大学院法学研究科 教授
	岡田　隆太朗	一般社団法人日本ディープラーニング協会 専務理事
	岡本　浩一郎	一般社団法人ソフトウェア協会 副会長／株式会社リアルソリューションズ 代表取締役社長
	工藤　郁子	大阪大学社会技術共創研究センター 特任准教授
	殿村　桂司	長島・大野・常松法律事務所 弁護士
	中尾　悠里	富士通株式会社富士通研究所人工知能研究所 プリンシパルリサーチャー
	永沼　美保	一般社団法人日本経済団体連合会デジタルエコノミー推進委員会 国際戦略WG 主査／日本電気株式会社 品質・エンジニアリング推進部門 主席プロフェッショナル
	平野　晋	中央大学国際情報学部 教授・学部長
	福岡　真之介	西村あさひ法律事務所・外国法共同事業 弁護士
	松原　実穂子	日本電信電話株式会社 チーフ・サイバーセキュリティ・スト

ラテジスト

○ ヒアリング対象者

Eunice Huang　Google Asia Pacific　AI・新興技術政策部長

小島　治樹　日本マイクロソフト株式会社　政策渉外・法務本部　政策渉外ディレクター

小俣　栄一郎　Facebook Japan 合同会社 公共政策本部 部長

須藤　修　中央大学国際情報学部 教授・ELSI センター 所長

根本　宗記　日本電信電話株式会社 技術企画部門 AI ガバナンス室 室長

高松　英生　株式会社三井住友フィナンシャルグループ 常務執行役員

４.議　題　AIのリスクと制度的対応について（ヒアリング）

５.資　料

資料１　Google Asia Pacific 発表資料

資料２　日本マイクロソフト株式会社 発表資料

資料３　Facebook Japan 合同会社 発表資料

資料４　中央大学国際情報学部 須藤修教授 発表資料

資料５　日本電信電話株式会社 発表資料

資料６　株式会社三井住友フィナンシャルグループ 発表資料

参考資料　AI 制度研究会 構成員名簿

６.議事要旨

○ ヒアリングに先立ち、松尾座長より以下の挨拶があった。

・ 先月２日、岸田総理、高市大臣御出席の下、ＡＩ戦略会議と合同で第１回研究会を開催した。構成員からは、生成ＡＩの登場によってリスクは大きく変わっている、ガイドラインはアジャイルでよいが守らない者もいる、国際整合性が必要、制度と技術の両面からＡＩの安全性を高めるべき、といった様々な意見を頂いた。

・ これを受け総理からは、イノベーション促進とリスク対応の両立、変化の速さへの対応、

国際的な相互運用性、政府による適切な調達、利用の４点を原則として検討を進めるようにと指示があった。

○ Google Asia PacificのEunice Huang様より発表と質疑応答があった。内容は以下のとおりである。

なお、質疑応答において、質問は「□」、回答は「■」とする

（発表）

・ 汎用AIは活用事例が一つではない以上、one size fits allのアプローチはない。
・ AI規制に対するアプローチとしては、コーディネート・整合性・焦点の三つの観点が重要。
・ ハブ＆スポークのガバナンス・アーキテクチャで各セクター毎に独自の専門知識を生かして規制すべき。政府はハブの役割を担うことにより見識を深めることが可能。
・ 基本は既存の法的枠組みを改善し、それでは対応不可のギャップを埋める事に重点を置く。
・ 国際的な枠組みや規格に整合している事が重要。G7の枠組みやISOを基礎とすることを提案する。
・ エンド・アプリケーションに焦点を当て、リスクベースのアプローチを推奨する。また、ユースケースのエンドポイント毎での規制が理想。

（質疑応答）

□責任を明確化するという点で、開発者の責任をどう考えるか。また、リスク判断のために開発者による情報開示があるが、利用者に対する責任はどう考えるか。

■ スクを最も理解できるプレーヤーが責任を持つべきである。例えば、開発者が持つ責任は、情報・文章を提供やモデルをどういう形で使うべきであるのかなどを推奨すべきである。Googleではテクニカルレポートにおいて既知のリスクについて情報提供している。また、開発者は利用者との接点がないので、利用者からのインシデント報告においては実装者が責任を持つべき。

□ ハブ・アンド・スポークのアプローチを取る際に、どういった形で政府の見識を深めることに役立つのか。

■ アメリカの場合はハブとしてNISTが各分野の事業者とコミュニケーションをとり、全体的なフレームワークを所有している。そのため、各セクターにおける知識がハブに積みあがっていくこととなる。

□ EEOCが既存の法律をAI使用時にも執行していくとイニシアチブを発表しているが、日本も

同様に積極的な既存の実定法の法執行をやるべきという意見で良いか。

■ 既存法でAIの新しいリスクに対応できると考える。強制執行能力があるならば、例外を増やす必要はない。また、既存の法律を検討した上で政府がそこにギャップがあると感じた場合に的を絞ってカバーすべきである。

□ 責任明確化のためにはある程度の情報開示が必要である。そのため、一定の守秘義務の下で情報を共有し合って、本当の責任者が分かるようにすることが必要だと思うが、どのようにお考えか。

■ 透明性は重要なので、情報開示には原則賛同するが、企業として共有すべき範囲や頻度などの論点があり、プライバシーや知的財産を踏まえて、透明性のバランスが重要である。

□ AIが使われる文脈においてリスクは変わる。国際的な相互運用性を保ちつつ、国によって文化的背景が異なるAIへのリスクに配慮していることはあるか。

■ ファインチューニングを行う際に文化的ニュアンスを反映しているが、安全性などの根本的な要素は国際規格に合わせる必要がある。そのため、国際的なベンチマークや誰でも活用できる標準的なアプローチが重要である。

□ AGIのような人間の知能を大きく超えるような人工知能に対して何らかの規制を課す、または将来的なリスクも含めて規制を課すという考えに対する意見を伺いたい。

■ 業界として積極的に研究を進めている分野であって、初期段階であると考える。現状、AGIがどのレベルで到達したと見なすかコンセンサスは無いため、まずはモデルの能力についてどのレベルに達しているかなどリサーチすべき。

□ 開発者側として、AIのテクニカルレポートという形で情報提供を行った際にレポート作成におけるコストや専門的な人材、プロトコルなど詳しく伺いたい。

■ 会社から多くのリソースを活用してテクニカルレポートを作成。また、開発者側からの情報提供に対し一貫性があると、実装者や利用者、政府等にメリットがあると考える。

○ 日本マイクロソフト株式会社の小島様より発表と質疑応答があった。内容は以下のとおりである。

（発表）

・ AI活用の最大化のためにガバナンスは必要。ガードレールがあるからこそ活用が最大化できる。

・ AIの規制を考えるときには、技術的なアーキテクチャに沿って制度は作られるべき。
・ アプリケーションのレイヤでは既存の法律を遵守する必要有。モデルやインフラのレイヤでは、グローバルガバナンスとして広範で包括的で、世界的に整合性のある制度を求める。
・ 非常にリスクの高いモデルは、政府が監視・管理する必要があるため、通知義務を課すべき。また、その仕組みを担保するため、インフラはライセンス制とすることを提案する。
・ 科学的根拠に基づいた基準が重要。アメリカではAIの規模をひとつの指標としているが、相互運用性の観点からこれに倣うのが良いのではと考える。
・ AISIや広島AIプロセスなどの国際相互運用性の高いものを基に、日本がリーダーシップを発揮していくことを期待する。

（質疑応答）

□ アプリケーションやモデル、インフラという3段のところで、ＡＩ規制がどのように相互に関連し合っているのか、それとも個別に運用されているのかを伺いたい。

■ モデルの開発者とアプリケーションの実装者では見える情報が異なるため、レイヤ間での情報共有が重要だと考える。また、それぞれが見える情報に応じて比例的に責任を負うことが大事。

□ 政府に対して情報を報告することで安全保障的な観点から安全性が確保されるということでよいか。米国の大統領令と関連して考えているのか。

■ 然り。通知に対応しない人もいるため、能力の高いモデルはインフラレイヤが把握のうえ、政府への通知を確認することで仕組みを担保することが重要。また、政府が情報を把握する必要性は、大統領令の考え方に沿ったものである。

□ 発表の中であった高度なモデルはどのような定義で考えられているか。また、ある閾値で規制等を設定すると逃れる可能性等を踏まえて伺いたい。

■ トレンドとしてスモールランゲージモデルがあり、将来的に必ずしも計算量が最適ではなくなる可能性はあるかもしれないが、現時点では計算量をベースに閾値を設定するのは米国や欧州との相互運用性の観点から良いと考える。また、モデルの中身でなく計算量で測れるのはインフラの事業者が閾値を超えたことを把握できるためリーズナブルである。

□ EUのAI Actのようなリスク評価等の義務を課したアプローチについてはどう考えるか。

■ スクベースアプローチに賛同したうえで、リスクがあるところに最小限の規制を課す観点から、義務として実施すべきセーフガードはどこまでの範囲とするのかという論点が重要である。

○ Facebook Japan合同会社の小俣様より発表と質疑応答があった。内容は以下のとおりである。

（発表）

・ 開発者・研究者が自由に改良・改変できるLlamaモデルを無料で提供し、AIの民主化・研究促進に貢献。

・ 何をリスクとして特定するのか、それをどう評価するか、がポイントとなる。一貫性のルール、国際的な相互運用性の確保が必要。

・ ルール形成においては、バリューチェーン全体を考え、各主体を責任主体として関与させるべき。リスクはユースケースによって異なる。もし提供者がモデルを不正に利用して損害を生じさせてしまっても、その下流の責任を上流に遡って課すべきではない。

・ リスクベースのアプローチが必要。オープンソースとしてリリースすることで、潜在的なリスクを「集合知」を使って洗い出せるし、対応をすることもできる。

・ リリースに当たっては当然、オープン化するメリットと潜在的リスクを考慮。社内外の専門家によるred teaming演習を繰り返し、様々なリスクへ対応した上で公開している。

・ オープンイノベーションでは透明性を確保しながら効果的に対応可能。AIを最大限に活用しつつ、リスクを最小限に抑えられる。

（質疑応答）

□ オープンソース全般に言える問題かもしれないが、オープンソースの中にバックドアが仕込まれるなどのリスクについてどう考えるか。

■ オープンソースの考え方は、そのような悪用する人間よりも、その悪用をブロックしようと貢献する人間が大半であるというものである。また、脆弱性が発見された場合にはそれを共有し、塞ぐような形のインセンティブを与えるなど、報奨金制度を設けてもいる。

□ 訓練データのポイズニングリスクについて、セキュリティ確保はどうしているか。

■ 例えば、アノテーションを含む開発段階から、特にプライバシーの観点には配慮している。また、prompt injectionといったセキュリティ・リスクに対しては、数次のred teaming演習はもちろん、Prompt Guardといった開発者ツールをオープンソースで公開するなどの対応している。

□ 御社のような影響力の大きい組織として、伝統的なデジュールとオープンソースにおける

優先順位やバランスをどのように考えているか。

■ 弊社はメタバース構築に向けて歩みを進めており、それを達成するためにAIは必須の要素である。巨額な投資を行って、オープンソースへの強いコミットメントを示している。またコミュニティと連携してリーダーシップを取る一方で、合意形成には時間がかかるかもしれない。多くの研究者の参加を募り、ロバストな基準を早く作ることが課題であると考えている。

□ 影響力のあるプロジェクトへの参加を重要視しているということか。

■ そのとおりである。弊社は、米国ではAI Safety Institute、またPAIなどを通じて、業界のスタンダード形成に貢献しており、日本企業が多く入るAI Allianceを含め、引き続き議論をリードしていきたい。

□ 法律では定義が非常に重要になるため、オープンソースの定義についてはどのように考えているかお聞かせ下さい。また、オープンソースの規制は緩くすべきと考えるか伺いたい。

■ 弊社として考える「オープン」の定義は、「モデルとその重みを責任を持って公開していること」と考えている。また、オープンソースはリスク対応の面で優れていることはインターネットの歴史が示している。漠然とした不安や憶測のみをもってAIの発展を阻害してはいけないと考える。

○ 中央大学国際情報学部の須藤様より発表と質疑応答があった。内容は以下のとおりである。

（発表）

・ OECD会議体の中で様々な議論がある。AIが引き起こした被害への明確な責任規定が必要。
・ AIライフサイクルにおける責任の在り方は開発者、実装者、利用者等によって変わるため綿密な設定や業界毎のガイドラインなどが必要。
・ AIを組み込んだDXは今後重要であり、マルチステークホルダーズの意見聴取を重視すべき。
・ 規制の在り方の検討においては、マルチモーダルAIの性能検定が必要。シミュレーションをした上で社会システムの設計を考えるべき。
・ 他国はAGI（汎用AI）に対する実施規則やそれに伴う意思表明など、政府の政策でも緻密に考えられているため、日本も同様に検討すべき。

（質疑応答）

□ AGIについては空想の世界ではなく、もう対応すべきことか伺いたい。

■ 生成AIだけではハルシネーションの問題もあり、まだかと思うが、EUや様々な国際組織では既に議論がされているため、日本でも考えないといけない段階である。

○ 日本電信電話株式会社の根本様より発表と質疑応答があった。内容は以下のとおりである。

（発表）

・ AIガバナンス室を設置し、リスクベースのアプローチを実施。
・ AI事業者の観点（ビジネス・リーガル・レピュテーション）でリスクを評価。
・ 全体のグランドデザインとしては、イノベーション阻害のリスクがあるため、基本はソフトローとし、規制は必要な範囲で留める。
・ 予防の側面があるホワイトリスト型の場合は過度な事実上の規制となる懸念が高いため、法規制は事例を元にしたブラックリスト型が望ましい。
・ 自国産業保護の観点からジャパンファーストなAI施策を期待する。

（質疑応答）

□ AIサービス・モデルの評価・認証について詳しく伺いたい。

■ 外部サービス使用時に学習データに不適切なものは使わないなどといったことを、一企業の取組として示している。また、認証型にすると基準設定が用途によって変化するため、合理的な設定は難しい。

□ 被害事例を元にした制度整備のために、公開されていない情報への対応はどう考えるか。

■ ソフトローで運用しつつ、その中で生じた被害事例を集めてハードロー化することが大事。そのためにはデジタルプラットフォームによって権利者から集めていくことが重要。

□ 制度的には著作権法のような法律での対応を念頭に置いているのか、あるいはAI法の透明性義務や権利侵害抑止義務のような対応を念頭に置いているのか。

■ 海賊版の情報を元に学習されているのではないかというAIが出ていると認識している。しっかりとソフトローやハードロー作っていくことも大事だが、日本の法律だけでは対応が難しい見方があるため、しっかりと国際社会に訴えて連携することも必要である。

□ 法律に関して、いわゆる強制力が必要だという考えか。

■ ソフトローで方向付けを行った上で、実際の被害事例をもとに必要なところにハードローを整備。予防的措置は行き過ぎるとイノベーションを阻害する。

□ 被害事例ベースでの法整備イメージを伺いたい。

■ 被害事例を分析した上でモデル・ユースケース毎など止めるべきポイントの検討が必要。

□ 透明性のためには、情報開示を推奨や強制で実施する方法が挙げられるが、まずはガイドラインから始めるべきという考えか。

■ まずはガイドラインベースで示し、その中で生じた被害事例をもとにハードローにしていくべきだと考える。

○ 株式会社三井住友フィナンシャルグループの高松様より発表と質疑応答があった。内容は以下のとおりである。

（発表）

・ SMBCグループのAIの利活用に関する社内ルールは、グローバルでの規制・ガイドラインを参考に整備し、AIを新規に活用する場合は社内の関連部門が審査している。

・ 各地域で規制が異なると各法令への対応負担が大きくなってしまうため、もし日本で法令が制定される場合には、各国の内容とも整合性を取った形での制定を希望する。

・ システムにおける要件や方法を考える「AI企画者」も一つの主体として入れてはどうか。開発するAIの用途やリスクといった内容は、上流工程の設計で検討されるため、この企画者が何をすべきかを定めることで安全・安心なAI開発や利用を進むと考える。

（質疑応答）

□ 技術的な視点と監査の視点がある中で、リスク審査はどこが担当しているか。

■ IT部門がイニシアチブをとって審査を行っており、コンプラ面やリーガル面など様々なリスクがあるため、関係部門で各々が担当する箇所をチェックする体制を取っている。

□ AI企画者の提案は良い考えではあるが、AI企画者とAI提供者を分ける理由を伺いたい。

■ AI企画者とは既存のガイドラインでいうとAI利用者の一部ではあるが、AI活用の上流工程という大切な役割を担うAI企画者をAI利用者とは別に定義すべきと考える。

□ AIの活用、あるいはリスクを軽減していくために金融の業法として何か修正したほうがよいという意見はあったりするか。

■ 業法面で思い当たる点は、現状特段ない。

□ 金融業に限らず、ルールを定めてほしい所やAIを使われる立場として意見を伺いたい。

■ 悪用事例の収集と、その事例を元にAIが悪用されている部分の法整備が必要だと考える。

○ 本日のヒアリング全体を通して、各構成員よりコメントがあった。内容は以下のとおりである。

・ バリューチェーンの中で開発、提供、そして利用という人たちの間での責任の配分が重要
・ ライアビリティ側のルールが設計されていくと大きなリスク発生を現象させていく柔軟なインセンティブを与える可能性があるため、ライアビリティ側の議論を進めるべき。
・ AI のライフサイクルや物理インフラを含めたバリューチェーンについて責任分解に関してトータルに考えたアプローチが重要。
・ 被害対応・情報収集、更には定量・定性など指標や基準等の在り方を見直すべき。
・ ユースケース毎でのリスクで既存の法規制で対応すべきである点と共通するリスクをしっかりと定めるべき。
・ 安全性の確保には透明性の議論がある。情報開示をソフトロー/ハードローで行くのか、国際基準を待つか/今何かする必要があるかなど、日本としてどうするか検討が必要。
・ 業界やユースケース毎で実施し、エアポケットを拾っていく考え方に傾いている。
・ 標準化された手順やプロトコル・体制など大企業だからできる部分はあるが、それを様々な企業でも対応できるように認証的な形で落とし込む事が重要。
・ 専門家がいなくても、うまく機能するようなプロトコルや認証制度を作ることが必要。
・ 業界ごとにリスクの種類が異なり、ユースケース毎に検討する事が必要。
・ 企業では既に事業者ガイドラインを取り入れている動きがある。その上で、適宜、改訂、アップデートが必要なところはフトロー（ガイドライン）での対応となるのではないか。一方で、業種毎にリスクが異なるとしても、悪用事例の部分を牽制するような活動は共有することが必要。
・ ソフトローだけでは実行力が足りない。監督官庁がしっかりとモニターし、ガイドライン改正や省令、事業法改正など必要であれば対応すべき。
・ ガイドラインからモニタリング、法改正など段階を上げていく事も考えられる。
・ OSS は AI の民主化において重要であり、不要な規制は避けるべきだが、OSS を例外にするとそこが穴になる可能性がある。また、OSS 以外の事業者のみ規制があると不公平が生じる可能性もあり、OSS の取り扱いにも十分な議論が必要。
・ 現在の複雑な状況にある中で、AI の制度は、文理一体となり知識をもって熟考すべき。
・ 事例を集めた上でそれをベースに検討すべき。小さな企業でも対応可能であるべき。

- このAI制度研究会を通じて責任のあるAIの開発・利用のユースケースが集まってきたが、一方で、責任のある開発・利用をしていないところからの被害や危険性についての情報も集めるべきではないか。
- 認証、罰則、ブラックリストなど色々なやり方があるが、事例集めが一番難しい。
- 国際的な相互認証性が重要視され、標準化が政策に結び付く。
- 国際的な標準に合わせる、国策でやるという2つの観点を表裏一体で考えるべき。
- 事例が重要であり、その情報収集から始まり、ソフトロー、ハードローといった形で徐々に仕組みを構築していくことが必要である。
- 国際的な差異含め、日本としてどのような戦略をとっていくべきか考える必要。
- 今回のヒアリングで、開発者や利用者、国内外企業、法制度の専門家など非常に多くの方からお話を頂いて、大変多くの示唆を受けた。これらの意見を取りまとめて次回以降の場で議論頂きたいと考えている。

以上