「責任あるAIの推進のための法的ガバナンスに関する素案」に対するBSA | The Software Allianceからの提言

2024年4月11日

総論

BSA｜The Software Alliance[1]（BSA | ザ・ソフトウェア・アライアンス、以下、BSA）は、自由民主党デジタル社会推進本部の「AIの進化と実装に関するプロジェクトチーム」（以下、AI PT）のリーダーシップのもとで、責任あるAIの推進に向けた取り組みが進んでいることを高く評価しています。また、AIPTのWG有志から示された「責任あるAIの推進のための法的ガバナンスに関する素案」（以下、素案）の目的を我々は支持しています。素案においては、基本的人権をはじめとする国民の権利利益が侵害されるリスクを最小化しつつ、AIの健全な発展による利益を最大化することが目指されています。BSAとその会員企業は、これらの目標を達成するために、AI PTに協力していきたいと考えています。提示されている共同規制モデルにおいては、民間分野の関与が明示されていることから、素案の具体化に向けて、今後、建設的な意見交換をしていけることを我々は期待しています。

BSAは、世界のソフトウェア産業を代表する主唱者です。BSAの会員企業は、AIを含む最先端のサービス開発の第一線におり、その製品は経済のあらゆる分野で企業に利用されています。[2] 例えば、BSA会員は、クラウドストレージやデータ処理サービス、顧客関係管理ソフトウェア、人事管理プログラム、ID管理サービス、サイバーセキュリティサービス、コラボレーションソフトウェアやシステムなどのツールを提供しています。そのため、デジタルトランスフォーメーション（DX）を促進するテク

[1] BSAの活動には、BSA’s members include: Adobe, Alteryx, Altium, Amazon Web Services, Asana, Atlassian, Autodesk, Bentley Systems, Box, Cisco, Cloudflare, CNC/Mastercam, Dassault, Databricks, DocuSign, Dropbox, Elastic, ESTECO SpA, Graphisoft, Hubspot, IBM, Informatica, Kyndryl, MathWorks, Microsoft, Nikon, Okta, Oracle, PagerDuty, Palo Alto Networks, Prokon, Rockwell, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, Workday, Zendesk, and Zoom Video Communications, Inc.が加盟企業として参加しています。詳しくはウェブサイト（http://bsa.or.jp）をご覧ください。

[2] “Artificial Intelligence in Every Sector（あらゆる分野における人口知能（AI））” https://www.bsa.org/files/policy-filings/06132022bsaaieverysector.pdf（英文）

22F Shibuya Mark City
1-12-1 Dogenzaka, Shibuyaku,
Tokyo 150-0043

P +81 3 4360 5473
F +81 3 4360 5301
W bsa.org

Japan Representative Office

ノロジーの大きな可能性や、責任ある AI を最も効果的に支える政策について、独自の見識を持っています。

本提言においては、AI に関する法的枠組みの策定を進めるにあたり、BSA と BSA 会員企業、また、その他の利害関係者を含めることを、AI PT に推奨します。下記に詳述する BSA の提言内容は以下となります。

- AI ガバナンスにおいては、一貫性のある、国際的に認知されたアプローチを採用すること。
- 法的枠組における不整合を避けること。
- リスクの高い AI システムの利用に焦点を当てたリスクベースのアプローチを採用すること。
- AI の開発・導入における役割に基づき、適切な主体に責任を割り当てること。
- 公平で相応な義務を割り当て、過度に規定的で負荷の高い要件を課するのを避けること。
- 立法・規制プロセスにおいてマルチステークホルダーの関与を促進すること。
- 第三者による検証を優先もしくは要件として示唆するのを避けること。

世界中の政策立案者が AI に対する規制アプローチに取り組む中、今日のテクノロジーをとりまくエコシステムは、そのグローバルな性質により、イノベーションを促進するための協調的な政策対応を必要としています。我々は、各国が協力してマルチステークホルダー間の対話を促進し、リスクベースの政策アプローチに関する共有できるビジョンを策定し、共通の AI 課題への対処と責任ある AI ガバナンスに関する規範を前進させることを推奨します（例えば、規制に対するリスクベースのアプローチや、AI バリューチェーンにおける公平な責任分担）。また、グローバルパートナー間で、共通の AI 用語と分類体系について合意することも必要です。これにより、革新的事業者が、自信を持って柔軟に有益なアプリケーションに技術を採用することが可能となります。このような調和的なアプローチを素案において反映することを推奨します。AI はグローバルな技術であり、国境を越えて開発・利用される、多くの国際協力の成果です。国境を越えるリスクに対処し、技術の開発と利用に関する国際協力を可能にするためには、世界的に一貫性のある AI ガバナンスの枠組みを定めることが重要です。ガバナンスの枠組みは、国際的なベストプラクティスに沿い、世界的な枠組みとの相互運用性を優先すべきです。

リスクの高い利用に焦点を当て、グローバルな相互運用性を確保する

規制を考える上では閾値が重要となることから、我々は、「AIシステム」の定義を国際的に認知された定義と明確に整合させることを奨めます。具体的には、以下のOECDのAIシステムの定義を採用することを推奨します。

「AI システムは、明示的又は暗黙的な目的のために推測するマシンベースのシステムである。受け取った入力から、物理環境又は仮想環境に影響を与える可能性のある予測、コンテンツ、推奨、意思決定等の出力を生成する。AI システムが異なれば、導入後の自律性及び適応性のレベルも異なる」[3]

AIシステムが国際的な文脈で開発され、展開されることを考えると、AI技術のさらなる普及と利用を促進するためには、AIに適用される定義が、異なる法域にまたがって運用可能とすべきです。OECDの定義のように、国際的に認知されたAIシステムの定義を用いることで、日本の政策の国際的な整合性が保たれ、本AI法案に関する議論、採用、また遵守を促進することが可能となります。

BSAは、リスクの高いAIの利用にガードレール（安全対策や基準）を設ける規制の枠組みを支持します。AIは様々な文脈で利用される可能性があるため、政策立案者が焦点を当てるべきなのは、消費者に最も大きいリスクをもたらす利用です。AIシステムは、ビデオ通話の背景雑音の検出・低減から製造工程の最適化まで、大きなリスクを伴わない、幅広い場面で利用されています。文書で使うフォントの種類を予測するのに利用されるAIシステムのような、低リスクのシステムについて、追加の義務を課す必要はありません。しかし、高リスクのシステムについては、開発者と導入者はリスクを評価し、軽減するための対策を実施する必要があります。

素案では、基盤モデルまたはフロンティアモデルに焦点を当てようとしていますが、これらの概念を定義して、新たな義務を課すことには慎重であるべきです。これらのモデルの能力や基盤技術は進化し続けており、最も重大な危害をもたらす利用に対処するためには、現時点ではリスクベースのアプローチの方が適しています。さらに、これらの課題の検討を続けていく上では、政策の不要な分断を最小化し、国際的な相互運用性を促進するために、定義や規制のアプローチを可能な限り他国と一致させることが重要となります。これにより、AI開発者・導入者、規制当局、消費者が、リスクを特定して最小限に抑えながら、頻繁に国際的に提供されるこの技術から得られるメリットを最大化することが可能となります。また、これにより、AIの法整備に向けた、世界の他の地域における広範な取り組みや、この分野における先駆的リーダーシップの恩恵を受けながら、そして、これらを基礎として、日本の制度を策定することが可能となります。どのような法的枠組みであっても、技術の進歩に合わせて概念や定義を更新できるように、柔軟性と適応性を確保することが重要です。

以下に詳しく説明するように、我々は、AIの開発者や導入者が、リスク管理プログラム、影響評価、リスクの高いAIの利用に関する内部検証を実施することを支持しています。AI開発者（基盤モデルの開発者を含む）は、AIのバリューチェーンに沿って、また、関連するリスクのレベルに基づいて、モデルの能力、限界、検証、セキュリティに関する情報を提供すべきです。これにより、AIエコシステムにおける導入者やそ

[3] OECDによるAIシステムの定義の更新に関する説明（2023年11月29日）
https://oecd.ai/en/wonk/ai-system-definition-update

の他の主体が、AIソリューションの特定の高リスク利用において生じ得る問題をよりよく理解し、特定し、対処することが可能となります。そして、この急速に変化し、重大な影響を及ぼす技術と規制に関して、更なるコンセンサスが形成される中、日本のアプローチが十分に反映されるよう、日本がAIガバナンスの国際的な議論に引き続き積極的に参加することを奨めます。

法的不整合を避ける

技術的中立性の原則を無視し、既存の法令との法的不整合を招きかねない法案は避けるべきです。どのような法案においても、焦点をあてるべきなのは、規制の隙間を埋めることです。多くの AI システムはすでに既存の法律により規制されています。いかなる AI 法制もこのことを念頭に置き、既存の法的要件の対象となっている AI システム事業者に対し、重複的で、不整合で、不要な要件を新たに課すことは避けるべきです。

リスクベースアプローチの採用

個人に対して高いリスクをもたらすユースケースに焦点を当てた、リスクベースの AI 政策へのアプローチを我々は強く支持しています。AI の高リスクな利用には、例えば、住宅、雇用、与信枠、教育、公共の場へのアクセス、医療、保険に関して、個人の適格性を判断し、これらの提供・拒否につながるものが含まれます。素案には、特にリスクの高い領域における AI システムについて安全性検証を行うことが記されていますが、我々は、素案における要件を「領域」ではなく、高リスクのユースケースに限定することを推奨します。AI の多様な利用によってもたらされる便益、損害、政策的考慮事項は大きく異なります。

例えば、AI システムは、医療提供者のスケジュール管理を容易にしたり、請求の問題に対処したり、あるいは日常的な管理業務を助けるために利用されるかもしれません。これらは通常、医療分野における AI の「低リスク」な利用となります。一方、健康保険の還付や特定の治療の適格性の判断に関わる AI システムは「高リスク」な利用と考えられるかもしれません。このことから、低リスクのユースケースに遵守義務を課すことに我々は反対します。これらのユースケースには、ビデオ通話での背景ぼかし、自動修正、電子メールのスパムフィルター、ウェブ検索エンジン、テレビ番組のレコメンデーションなども含まれます。このような低リスクの技術に遵守義務を課すことは、事業活動を大幅に停滞させることになりかねず、また、一般に受け入れられ、広く利用されている技術を使って事業活動が行われることを期待している消費者に有意義な恩恵がもたらされない可能性があります。

AIエコシステムにおけるAI主体間の均等のとれた責任分担の確保

AIシステムは極めて広範な状況で利用される可能性があり、AIシステムから生じるリスクは特定のユースケースによって大きく異なる可能性があります。素案においては、基盤モデルの開発者の遵守要件に焦点が当てられていますが、これでは立法目的を達成するには不十分かもしれません。AIのバリューチェーンは多様かつ複雑であるため、我々は、遵守するのに最も適した主体に責任を割り当てることを推奨します。

開発者はAIシステムを設計、コーディング、製造する主体であり、導入者はAIシステムを利用する主体であるため、これら二つの組織は潜在的なリスクを特定し軽減する上で異なる役割を持つことになります。さらに、この二つのタイプの組織は、異なる種類の情報にアクセスすることができ、潜在的なリスクを軽減するために異なる措置を講じることができます。例えば、AIシステムを設計する開発者は、AIシステムの学習に使用されるデータの種類、システムの既知の限界、および意図される利用例に関する情報にアクセスできる立場にあります。対照的に、AIシステムを利用する導入者は、消費者に影響を与えるそのシステムの具体的な利用方法に関する情報にアクセスできる立場にいます。AIの説明責任を支えることに重点を置いた政策は、このような異なる役割を反映し、それに応じて義務を割り当てるべきです。

組織は、他の役割を担うこともあります。既存のAIモデルを組織の製品やサービスに組み込む場合などです。このような組織に課される義務も同様に、AIシステムを組織の製品やサービスに組み込む際の役割を反映したものでなければなりません。

規程的な透明性・報告要件を避ける

素案では、指定を受けた特定AI基盤モデル開発者は、第三者による脆弱性検証やAIの能力と限界の公開など、七つの義務を実施する体制を構築することが求められると記されています。また、この義務に基づき、指定された開発者は、政府または第三者機関（AIセーフティ・インスティテュート）に対して、義務の遵守状況を定期的に報告する必要があるとされています。

我々は、これらの義務を通じて関連リスクに対するガードレールを設けようとする素案の目標を支持しますが、基盤モデルの規制がモデルのリスクと能力に見合ったものであることが重要であると考えます。そのため、基盤モデルの開発者は、AIバリューチェーンに沿って、モデルの能力、制限、検証、セキュリティに関する情報を、関連するリスクのレベルに応じて提供すべきです。リスクの高いAIシステムの利用については、安全性、セキュリティ、正確性、有害なバイアスについて、着実な検証と評価を行うことを奨励します。一方で、AIの検証に関する既存の技術標準が発展途上であることを理解することも重要です。標準開発は、長年、自主的に、市場主導で、コンセンサスに基づくアプローチで開発されてきました。AIの検証に関しても同様に開発されるべきです。

我々は、AI が生成したコンテンツについて、電子透かしやその他の開示方法を利用することを支持します。しかし、広範な報告要件により、規制当局に審査書類が殺到し、また、企業が AI システムの開発・利用に関する専有（プロプライエタリ）情報や機密情報の開示を求められることを我々は懸念しています。

効果的な共同規制の促進

素案では、いわゆる「モニタリング・監督」の枠組みが共同規制の中核であると説明されています。共同規制は、強制的な介入措置に比べ、利害関係者間の迅速かつ柔軟な対応を促す、変化の激しい事業環境に適した規制手法と考えられています。一方で、共同規制を恣意的に適用すると、やりとり自体が過度な事務負担となり、事業者の日常業務を不用意に阻害するおそれがあります。そのため、共同規制の実施にあたっては、以下の点を考慮することを推奨します。

- 一般市民への説明や各企業の自主性を尊重するために、行政の関与が必要かどうかを慎重に検証すること。特定の AI システムの開発や導入リスクを効果的にコントロールすることにならない取り組みの実施を企業に求めないこと。
- 特定 AI 基盤モデル開発者への確認や照会に際しては、求める情報がどのように法の目的と関連しているのか、また、法の基本原則に合致しているのかを、監督当局が十分に説明すること。
- 企業秘密を含む可能性のある情報に関しては、特定 AI 基盤モデル開発者が差し控えるのを可能とすること。情報を求める監督当局は、提供された情報を機密として取り扱うこと。

第三者による外部安全性検証や脆弱性の検出・報告の強要を避けること

素案における七つの義務には、自社・外部による安全性検証の実施、第三者による脆弱性の検出と報告が含まれています。安全性の検証と検出がリスクを特定する上で重要であることに我々は同意します。しかし、組織が常に外部による検証を実施すべきであると示唆することには反対します。組織が外部による検証の実施を選択する状況もありますが、自社検証（AI システムの開発担当チームではない、独立した従業員チームによる実施が可能）は、企業秘密、情報やネットワークのセキュリティを危険にさらすかもしれない情報、また外部検証で発生する専有情報の共有といった懸念を生じさせることなく、リスクを特定し、軽減することを可能とします。このため、義務の焦点は自社検証にあて、独立した外部による検証を外すことを推奨します。

また、素案では、外部監査を示唆する、第三者による脆弱性の検出と報告も含まれています。現時点では AI に関する監査可能な基準が成熟していないため、外部監査の

活用には慎重であるべきと考えます。現在、以下のいずれかを企業が実施する上で、既存の手順やベストプラクティスはほぼありません。

（１） AI システムを監査できる信頼できる法人を選択する。

（２）上記の監査法人がどのような基準を適用すべきかを決定する。

ISO はいくつかの AI 関連規格を発行していますが、多くの規格はまだ開発中です。また、現在、AI システムに対応する十分な自主的コンセンサスに基づく規格が不足しています。共通の基準がなければ、監査の質は大きく異なります。監査によって異なる基準で測られる可能性があり、客観的な基準に基づく評価を得るという目標が損なわれます。

また、BSA は透明性を促進する必要性を理解していますが、機密情報や専有情報を含む監査結果を公表することを事業者に求めないことを推奨します。公表することは、AI システムの厳格な評価を受ける意欲を企業に失わせることになります。このような理由から、外部監査は AI ガバナンスを達成するための適切な解決策ではないため、この義務を外すことを奨めます。

結論

BSAと会員企業は、AIに対する効果的なセーフガードを策定するというAI PTの目標を支持しています。本提言を共有することに加え、素案の意図をより深く理解し、この取り組みを今後どのように支援していけるかについて話し合う機会を頂ければ幸いです。