AI 新時代におけるAI国家戦略に関するBSA| The Software Allianceからの提言

2023年5月5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下に発足した「AIの進化と実装に関するプロジェクトチーム」（以下、プロジェクトチーム）が主導した「AIホワイトペーパー－AI 新時代における日本の国家戦略－」（以下、ホワイトペーパー）において、AIのさらなる導入が奨励されたことを高く評価します。また、AI開発能力の育成・強化や官民におけるAI活用の推進・支援に焦点を当てた新たなAI国家戦略を日本が策定する必要がある、とホワイトペーパーで提唱されたことも我々は歓迎します。プロジェクトチームが、社会の生産性、品質、効率性を向上させる上で、AIがもたらす利益を十分に認識していることを我々は心強く思います。 経済産業省が策定した「AI 原則実践のためのガバナンス・ガイドライン」[2]で示されたように、産業界の自主的な取り組みを支援する日本のアプローチをBSAは高く評価し、支持します。 BSAとその会員企業は、日本の経済成長、競争力、雇用創出を支援するために、AIが責任を持って開発・利用されるよう、プロジェクトチームおよび日本政府と協力していけることを期待しています。

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、クラウドコンピューティング、データアナリティクス、人工知能（AI）など、世界的な経済成長を牽引する、ソフトウェアによって実現されるイノベーションの最前線にいます。社会のあらゆる分野でAIのメリットが活用され、人々の生活が劇的に変化し、複雑な問題が解決されることをBSAの会員企業は支えており、世界中のイノベーションを促進しています。[3] BSAの会員企業は最先端技術の開発を主導しており、そのた

[1] BSAの活動にはAdobe, Alteryx, Altium, Amazon Web Services, Atlassian, Autodesk, Bentley Systems, Box, Cisco, Cloudflare, CNC/Mastercam, Dassault, Databricks, DocuSign, Dropbox, Graphisoft, IBM, Informatica, Juniper Networks, Kyndryl, MathWorks, Microsoft, Nikon, Okta, Oracle, Prokon, PTC, Rockwell, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, Unity Technologies, Inc., 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] https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20220128_1.pdf

[3] 詳細については、2022年6月13日にリリースされた、“Artificial Intelligence in Every Sector”を参照ください。 https://www.bsa.org/policy-filings/artificial-intelligence-in-every-sector

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

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

Japan Representative Office

め、BSAは、先端技術の大きな可能性、また、それらの責任ある開発と利用を支援するための最良の政策、その両方に関して、独自の見解を持っています。

AIの導入は、組織、消費者、社会に間違いなく利益をもたらします。一方、この技術が責任をもって開発・導入されなければ、重大なリスクをもたらす可能性があることも我々は認識しています。また、BSAは、AIが有害な方法で使用される可能性があることも理解しています。例えば、AI システムは、個人を違法に差別するかもしれません。そのため、そのようなシステムが、意図しない偏見に関するリスクを特定し、軽減するために徹底的に検証されていることが、一般市民に保証されるべきです。この点で、我々は、AIの利点を促進する一方で、リスクの高いAIからの保護策を再検証し、AIに対する国民の信頼を築いていくことを提言しているホワイトペーパーを支持します。

この目的の達成を支えるために、BSAから以下の見解と提言を述べさせて頂きます。加えて、ホワイトペーパーの提言内容の実現に役立つと思われる関連資料を添付します。添付資料に含まれるのは、AIバイアスリスクを特定して軽減する影響評価の初めてのフレームワークである「バイアスに挑む：AIの信頼性構築に向けたBSAのフレームワーク」[4]（以下、BSAのフレームワーク）、また、開発者と導入者の異なる役割を説明した「AI 開発者と AI 導入者：重要な違い」[5]です。これは、AI市場における組織の役割に応じた義務を考慮するための資料です。

AI製品やサービスは、これらの技術に対する社会の信頼と信用構築によって、うまく採用されるべきであるという、プロジェクトチームの意見に我々は同意します。その信頼を得るために、BSAは、AIを開発・利用する組織に対し、その技術の利用がもたらす固有の機会及びリスクについて説明するよう促しています。また、政策立案者は、責任あるイノベーションを支援する法的・規制的な環境を整備することで、AIに対する国民の信頼と信用を高めることができます。

そのため、BSAは、AIに関連するあらゆる規制に関し、以下を推奨しています：

- リスクの高いAIシステムのみに適用する
- 規範的な適合性評価要件は避ける
- 影響評価の採用を支持する
- AIの開発者と導入者の異なる役割と責任を認識する
- 国際的に認知されつつある規格と整合させる
- AIの責任ある開発・導入に対応するためのライフサイクル・アプローチを取り入れる
- イノベーションを促進するデータおよび知的財産の政策を維持する

[4] https://bsa.or.jp/wp-content/uploads/2021bsaaibiasjp.pdf
[5] https://www.bsa.org/files/policy-filings/jp04102023aidevdep.pdf

規制はリスクの高いAIシステムのみに適用すべき

ホワイトペーパーでは、AI規制に関する新たなアプローチにおいて着目すべきリスク領域として以下の三つを挙げています：１）重大な人権侵害、２）安全保障、３）民主主義プロセスへの介入。我々は、検討されている規制の範囲をリスクの高いAI利用に限定するアプローチを支持します。AIのエコシステムは幅広く、多様な技術、ユースケース、幅広いステークホルダーを包含しています。AIのリスクは本質的にユースケース固有のものであるため、いかなる規制も、一般市民に高いリスクをもたらす技術の特定の適用に焦点を当てるべきです。[6]また、特定のユースケースに関連する固有の考慮事項を説明する上でも、柔軟性を備えた規制である必要があります。リスクの高いユースケースを定義し、リスク判断においては、分野ごとのアプローチを避けることが重要です。例えば、給与管理AIや文書作成に使用されるAIシステム（例えば、同義語や構文を提案するワープロソフトのサポート）は、たとえ国家安全保障の文脈で使用されても、本質的にリスクが高いとは言えません。

また、ホワイトペーパーで特定された三つのリスク（重大な人権侵害、安全保障、民主主義プロセスへの介入）に対処するためにAIを規制する前に、これらのリスクを軽減するためにAIを最大限活用することが重要です。例えば、AIの技術により、ディープフェイクやサイバー攻撃の検出、また、サイバーセキュリティ、プライバシー、児童保護を強化することができます。[7]

規範的な適合性評価要件の回避

上記で強調したように、AIがもたらすリスクとそのリスクを軽減するための適切な仕組みは、状況に大きく左右されます。学習データの収集と利用、記録の保存、透明性、正確性、人間の監視に関する適切な仕組みは、AIシステムの性質とそれが導入される環境に応じて異なります。規範的なアプローチは、政策立案者や政府が防ごうとするリスクそのものに対処する取り組みを阻み、不要なコストを追加し、極めて複雑なコンプライアンスチェックを要します。規制において焦点をあてるべきなのは、ステークホルダーが自身のユースケースにどのような基準が適切かを判断する上で考慮すべき要素です。規制当局は、柔軟性に欠けるアプローチを避け、代わりにリスクベースの評価を促進する、プロセスや成果指向の政策解決策に焦点を当てるべきです。

---

[6] 高リスクには、重大な決定（consequential decision）が含まれるAIのユースケースが含まれます。つまり個人に対して法的または類似の重大な影響を及ぼすようなAI導入者による決定です。例えば、住宅、雇用、信用、教育、公共の場へのアクセス、医療、保険などに対する個人の資格・適格性を判断し、その提供または拒否につながるような決定です。

[7] 例えば、サイバーセキュリティ企業がAIを利用し、IPアドレスの位置などの情報に基づいて、ログインを企てる行為にリスクスコアを割り当て、悪意のある行為者が他のユーザーの銀行口座にログインを企てているかどうかを判断し、ログインしようとする行為を拒否することが可能となります。また、機械学習とパターンマッチングを利用することで、クラウドインフラ上でホストされている機密データを発見し、プライバシー保護を強化することができます。 同様に、機械学習モデルは、ディープフェイクを含む、改ざんされたメディアを検出するために利用することができます。また、AIツールを利用して、オンライン上における児童への性的虐待や性的搾取に対処する法執行機関を支援することができます。このような利用例は、AIがもたらす利点を示すとともに、個人を保護するためにAIを利用する、多様で積極的な活用を検討する必要性があることを明らかにしています。

また、AIシステムの市場導入前の適合性評価を確立することは避けるべきです。そのような義務は、不当な市場参入障壁となる可能性があります。むしろ、AIシステムの公正さを実現するために開発者や導入者が従うべき広範な目的とプロセスを特定する、ガバナンスに基づいた自己評価のアプローチの方が効果的です。多くのグローバルなAI開発・導入企業は、AI技術が安全かつ責任を持って構築・利用されることを保証するために、自主的な措置を講じ、AI倫理原則や企業の体制の中に正式な評価プロセスを確立しています。BSAフレームワークは、産業界のステークホルダーが一丸となって、AIリスクを特定し対処するための方法論を構築した一例です。

影響評価の採用を支持する

AI規制において、我々は影響評価の採用を奨励します。AIの責任ある利用を促進する重要な保護措置の一つが、リスクの高いAIシステムを開発または利用する企業が、影響評価と設計評価を実施するための包括的なアプローチを確立することを確実にすることです。影響評価は、環境保護から個人情報保護に至るまで、様々な分野で広く利用されており、システムが潜在的なリスクを考慮した形で設計されていることを示すことで、信頼を促進する説明責任の仕組みとなっています。

BSAは、重大な決定を行うために利用されるAIシステムについて、企業に影響評価と設計評価の実施を求めることを支持しています。このような評価や査定は、企業がAIリスクを特定し、文書化し、軽減するのに役立つ重要な説明責任ツールです。特に、違法な差別につながりかねない潜在的なバイアスを検出し、軽減する上でも有用です。影響評価と設計評価を義務付ける法律は、リスクの高い利用に適用され、開発者と導入者の要件を明確に区別する必要があります。

AIの開発者と導入者の異なる役割と責任を認識する

AIのリスク管理における特定の状況下では、様々な度合いの責任を負う可能性がある二つの重要な主体がいます：

- 開発者：AI開発者は、AIシステムを設計、コーディング、または製造する組織

- 導入者： AI導入者は、AIシステムを採用し、利用する組織　(ある事業者が自社で利用するAIシステムを開発する場合は、AI開発者とAI導入者の両方を兼ねうる)

政策や規制においては、このような区別を認識し、企業が契約によってリスク配分できるような柔軟性をもたらす必要があります。これらの異なる主体間のリスクの効果的な管理は、開発されるAIシステムの性質に依存します。開発者と導入者を区別することで、特定された義務がAIエコシステムにおける企業の役割を反映したものになります。開発者または導入者としての企業の役割に合わせて義務を設定することで、企業は対応する義務を果たし、消費者をより良く保護することが可能となります。例え

ば、開発者は、AIシステムの学習に使用するデータの出所を特定し、その特徴を説明することができますが、開発者は一般的に、他の企業がAIシステムを購入して導入した後に、そのAIシステムがどのように使用されているかの実態を見抜くことはできません。その代わり、AIシステムを利用する導入者は、一般的に、システムの利用方法、AIシステムからの出力、顧客からの苦情の性質、およびシステムの性能に影響を与えるその他の現実世界の要因を理解するために最も適した立場にあります。AIシステムが個人に対して生じさせる可能性のあるリスクプロファイルを理解する上で、導入者は、最も適した立場にいるのです。AI政策において、これらの異なる役割を反映した義務を策定することで、すべてのステークホルダーが、自身の組織がAIシステムにおける有害なバイアスをどのように特定し、対処すべきかを理解できるようになります。

国際的に認知されつつある規格と整合させる

プロジェクトチームと政府が、AI規制の新たなアプローチを模索する上では、それらが国際的に認知されつつある規格と整合していることを確認することが重要となります。この整合性により、国際的な相互運用性を向上させ、AIの開発者と導入者の双方を含む日本の組織が、利用可能な最先端のリソース、アイディア、選択肢から恩恵を受けることを促進することができます。国際標準化機構（ISO）のAI標準化委員会[8]は、AIシステムにおけるバイアスやAIにおける信頼性を高めるアプローチなど、10種類の標準化作業を完了しました。[9] 現在、ISO委員会はさらに27の規格を開発中です。国際的に認知された規格と整合性がとれない、あるいは、国際的に認知された規格より先行しすぎた国内規格を制定するリスクは、要件が新しい慣行と乖離し、日本でのAI開発を抑止し、技術が責任を持って開発・展開されることを保証する取り組みを阻害することになります。

また、AIシステムは国際的な文脈で開発・導入されるため、AIに適用される規制や基準は、AI技術のさらなる導入と利用を促進するために、異なる法域での運用を可能とすべきです。この点で、我々は、日本がOECDのAIの定義を採用することを提唱します。OECDは、人工知能に関する理事会勧告（Recommendation of the Council on Artificial Intelligence）[10]において、AIを「人間が定義した一定の目的のために、実環境あるいは仮想環境に影響を及ぼす予測、推薦又は意思決定を行う機械ベースのシステム」と定義し、AIシステムが「様々なレベルの自律性を備えて稼働するよう設計され

[8] ISO/IEC JTC 1/SC 42: https://www.iso.org/committee/6794475.html
[9] ISO/IEC TR 24027: 2021 (Bias in AI systems and AI aided decision making)
https://www.iso.org/standard/77607.html?browse=tc
ISO/IEC TR 24028:2020 (Overview of trustworthiness in artificial intelligence)
https://www.iso.org/standard/77608.html?browse=tc
[10] 2019年5月 人工知能に関する理事会勧告 https://www.soumu.go.jp/main_content/000642217.pdf
原文：https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
勧告では、AIのステークホルダーのコミュニティには「直接的なものであるか間接的なものであるかを問わず、AIシステムに関与するか、又はAIシステムから影響を受ける組織及び個人の全てが含まれる」としている。

ている」と明言しています。この定義は、欧州連合[11]を含む世界中の規制当局によって参照されています。 米国国立標準技術研究所（NIST）も、1月に発表したAIリスク管理フレームワーク[12]で使用するために、OECDの定義を適用しています。 OECDの定義のように、国際的に認知された定義を用いることで、国際的な整合、対話、採用、コンプライアンスを促進することができます。

また、AIやその他のツールの誤用を減らすために、業界団体が自主的に行っている取り組みもあります。例えば、多様なステークホルダーのコミュニティであるContent Authenticity Initiative（CAI）[13]は、コンテンツの真正性と来歴に関するオープンな業界標準の採用を推進しています。これにより、閲覧者は、撮影者、画像が生成された場所、ソフトウェアを使用して編集されたかなど、画像や動画の出所を知ることができます。これらの情報は、閲覧者がコンテンツの真正性を判断するのに役立ちます。900人以上のメンバーがいる同グループは現在、誤情報を防ぎ、AI利用に関する透明性を高めるためのオープンソース・ツールを開発しています。政策立案者や政府がこのような取り組みを支援することを我々は推奨します。

AIの責任ある開発・展開に対応するためのライフサイクル・アプローチを取り入れる

固定されたAIモデル評価では、AIシステムが現場に導入された際に発生する可能性のあるすべての問題を説明することはできません。例えば、バイアスは、システムのライフサイクルの複数の時点で、多くの異なるチャネルを通じて発生する可能性があります。例えば、モデルの学習に使用するデータ、システムが解決しようとする問題の定式化、あるいはモデルが意図した目的以外のシナリオで使用された場合などです。したがって、AIのリスク管理には、システムが意図したとおりに動作していることを確認するためのエンドユーザーによる継続的な監視を含む、ライフサイクル・アプローチが必要です。この問題に対処するため、我々はプロジェクトチームと政府に対し、AIのライフサイクルの設計、開発、導入の段階で、バイアスのリスクを軽減するために取ることができる措置を特定するBSAフレームワークを参照することを奨励します。

イノベーションを促進するデータおよび知的財産政策の維持

データ生産量の飛躍的な増加、リモートコンピューティング能力の向上、そして、より洗練されたアルゴリズムの開発が、AIの進歩に拍車をかけてきました。このような進展を活かし、AIの継続的な進歩を促進するためには、健全なデータ政策の環境が必要です。国際的なデータ移転は、予測モデルの開発からAIシステムの導入と利用に至るまで、AIライフサイクルの各段階で不可欠となります。AIシステムで使用されるデータは、地理的に分散した多くの情報源からであることが多く、越境データ移転の実

[11] 欧州連合のAI規制法の草案では現在、「AIシステム」を「... 人間が定義した一定の目的のために、コンテンツ、予測、推薦、または作用する環境に影響を与える決定などの出力を生成できるソフトウェア」と定義しています。
[12] 2023年1月26日　NIST AI Risk Management Framework https://www.nist.gov/itl/ai-risk-management-framework
[13] https://contentauthenticity.org/

現が不可欠となります。越境データ移転を不必要に制限する規則は、AIシステムが提供できる利益を例外なく制限することになります。日本は、イノベーションを促進するために政策を刷新する重要性を認識し、データが国境を越えて、途切れなく安全に行き来することを可能にするData Free Flow with Trust（DFFT）の概念を提案しました。[14]

先日のG7デジタル・技術大臣会合において、日本がリーダーシップを発揮し、G7参加国にデータ移転とAI活用に関する調和のとれた国際的な枠組みの構築を促したことを我々は高く評価します。特に、閣僚宣言[15]において、DFFT具体化に向けたパートナーシップのためのアレンジメント（Institutional Arrangement for Partnership、IAP）の設立が承認されたことを歓迎します。また、AIイノベーションのための環境をグローバルに実現するために、信頼できる AI のためのツール間のグローバルな相互運用性や、AI ガバナンスの相互運用性を促進するためのAIに関するアクションプラン[16]が承認されたことも歓迎します。我々は、これらの取り組みを支援していきたいと考えています。また、AIに関する新たな政策アプローチが、米国、EU、その他のG7参加国など、志を同じくする国々の立場と調和するようにすることを日本に求めます。

さらに、知識共有、連携、機械学習のような新技術の開発を向上させていくために、日本は、EUとともに、テキストマイニングとデータマイニングを可能とする著作権の例外規定を採用することで著作権法を刷新してきました。これにより、著作物に合法的にアクセスできるAI開発者は、公開されているコンテンツを使用してAIシステムを訓練し、無数の価値ある目的に使用できるデータの洞察を解き放つことが可能となります。これらは、投資とイノベーションを支援し、すべての人々にAIの恩恵をもたらすことを可能とする政策の例となります。

結論

BSAと会員企業は、効果的なAI政策の策定をしていく、という目標を支援するために、プロジェクトチームに協力していけることを期待しています。本提言を共有することに加え、ホワイトペーパー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継続的に頂ければ幸いです。

[14] “Data Free Flow with Trust (DFFT): Paths toward Free and Trusted Data Flows” https://jp.weforum.org/whitepapers/data-free-flow-with-trust-dfft-paths-towards-free-and-trusted-data-flows

[15] https://g7digital-tech-2023.go.jp/topics/pdf/pdf_20230430/ministerial_declaration_dtmm_jp.pdf

[16] https://g7digital-tech-2023.go.jp/topics/pdf/pdf_20230430/annex5_jp.pdf