

2023年7月20日

AIリスク対策を手がける米国発スタートアップ・Robust Intelligenceが 自民党および公明党の会議にて「第三者検証」によるAIガバナンスを提言

国内外の動向を踏まえ、今後のあるべき政策の方向性を提案

リスクを適切に管理したAIモデルの運用、“AI Integrity”の実現をめざすシリコンバレー発のスタートアップ・Robust Intelligence, Inc. (本社：米国カリフォルニア州、CEO：ヤロン・シンガー、Co-Founder：大柴 行人、以下：ロバストインテリジェンス)は、2023年7月11日(火)に開催された公明党「生成系AI活用検討委員会」、および2023年7月18日(火)に開催された自民党の第16回「AIの進化と実装に関するプロジェクトチーム」に登壇し、複雑化する生成AIリスクの捉え方と、その対策としての「第三者検証」について提言を行いました。



ChatGPTを始めとする生成AIの流行によって広がりつつある「AIリスク」への対応は、社会全体で急務となっています。日本政府もG7を経てこの問題への対処を急いでおり、ガイドライン改正等の新たな政策実装の作業を進めると共に、広島AIプロセス等を通じて国際的な議論を主導することを目指しています。

今回の自民党・公明党の会議も、行政および様々な関係者と議論して、政策的な方向性を見出すための場として開催されています。

ロバストインテリジェンスは、AIのリスク管理にテクノロジーの観点から取り組む日本市場でも珍しいプレイヤーとして、AIの開発段階から運用段階まで一貫してテストベースでリスク検証を行うプラットフォーム『Robust Intelligence Platform』を、諸業界の大手企業の皆様に提供しています (<https://www.robustintelligence.com/jp>)。このソリューションにより、客観的・自動的なAIモデルのリスク管理によってトラスト(信頼性)を確保する「第三者検証」を行うことで、大手企業が安心してAIモデルをデプロイできるようサポートします。

ロバストインテリジェンスでは、日本のAIガバナンスの行く末が多くの海外政府・企業からも注目されていることを踏まえ、IT/人材/金融等の諸業界の大手企業と協業してAIリスク管理のモデルケースづくりを進めると共に、最新のAIリスクに関する知見をもとにした政策への働きかけも積極的に行っています。

今回は、共同創業者の大柴および政策企画責任者の佐久間が、自民党・公明党それぞれの会議にて、生成AIのリスクの最新の研究動向とその捉え方を紹介するとともに、対策としての「第三者検証」を広めていくための枠組みを提言しました。概要は以下の通りです。

<講演概要>

■7月11日(火)公明党「生成系AI利活用検討委員会」

- ・講演者：政策企画責任者 佐久間弘明
- ・タイトル：「複雑化するAIリスクと企業のAIガバナンスの最前線」



■7月18日(火) 自民党 第16回「AIの進化と実装に関するプロジェクトチーム」

- ・講演者：共同創業者 大柴行人
- ・タイトル：「AIリスクに立ち向かう『第三者検証』～AIガバナンスを日本が主導するために～」
- ・資料：下記Webページに掲載予定

https://note.com/akihisa_shiozaki/n/n4c126c27fd3d



■講演内容(共通する部分が多いため、両講演を踏まえたサマリーを記載)

- 生成AIの普及により、AIリスクはかつてないほど複雑化・多様化している。こうしたリスクは幅広い社会に影響を与えるもので、個々の企業が単独で対応することが困難な場合も多く、サプライチェーン上の各プレイヤーが共同で問題に対処しなければならない

- そこで国内外の大企業が実践しようとしているのは、「第三者検証」とよばれるAIリスク管理のモデル。AIの開発・運用を行う主体と独立の主体が客観的・効率的にリスク検証することの重要性がわかってきており、諸外国の政策もこれを取り入れている
- 日本でこのモデルを定着させるには、リスク管理のあり方をマルチステイクホルダーで規定する認証制度のような枠組みが必要。政府やステイクホルダーと共同でその仕組みを作り上げ、G7・広島AIプロセス等で世界に発信する日本のガバナンスモデルとして定着させていきたい

ロバストインテリジェンスは今後も、AIリスクに対応する企業のガバナンス構築を支援し、日本市場におけるAI利活用の推進を図るべく、国に対する提言や、産業界を中心とした議論の場の構築といった活動を積極的に実施してまいります。

【Robust Intelligenceについて】

ロバストインテリジェンスは、2019年にハーバード大学の研究者たちによって創業されたシリコンバレー発のAIスタートアップです。

AIには、従来のソフトウェアには存在しない、精度劣化や倫理的問題などの特有のリスクが存在します。ロバストインテリジェンスはこうしたAIリスクに対応する企業のガバナンス構築を支援し、AIのリスクを適切に管理したモデルの運用“AI Integrity”の実現を目指しています。

具体的には、AIモデルの開発から運用に至るまでEnd to Endでリスクを防ぐため、AIモデルを自動で継続テストし、特定されたリスクに対してAIシステムを保護するプラットフォームを開発・提供しています。

リクルート、東京海上ホールディングス、SOMPOホールディングス、楽天グループ、Zホールディングス、セブン銀行、NTTデータ、NEC日本電気 (NEC Corporation)、ADP、Expedia、アメリカ国防総省など、国内大手企業や米国大手企業への導入実績を誇り、2021、2022年には「世界で最も有望AIスタートアップ100社『AI100』」に2年連続で選出されました。

【Robust Intelligence会社概要】

- ・設立年: 2019年
- ・所在地: US 555 19th Street San Francisco, CA 94107
- ・従業員数: 70人
- ・代表者: CEO & Co-Founder Yaron Singer、Co-Founder Kojin Oshiba (大柴 行人)
- ・主な投資家: Sequoia Capital、Tiger Global、Engineering Capital、Harpoon Ventures、In-Q-Tel
- ・URL: <https://www.robustintelligence.com/> (日本語版: <https://www.robustintelligence.com/jp>)