

CONCEPT

AIと共に生きる未来に向けて 「信頼」のためのAI倫理

世界経済フォーラム
AI（人工知能）・機械学習部門長

Kay Firth-Butterfield



日立製作所 執行役常務 CTO 兼 研究開発グループ長 兼
コーポレートベンチャリング室長

鈴木 教洋

近年、COVID-19によるパンデミックや気候変動など、地球規模で顕在化する社会の変化や課題に対応するべく、さまざまな分野でイノベーションが進展している。中でも大きなブレークスルーをもたらす可能性として期待されているのが、AIの活用である。

急速な進化を続けるAIが社会や暮らしの隅々にまで普及することによって、私たちの世界はどう変化するのか。AIがもたらすリスクを低減し、安心・安全を守りつつ活用するために欠かせない倫理、そしてそれを踏まえたアプローチとは――。

「信頼」というキーワードの下、新たなテクノロジーの導入に向けたガバナンス強化の観点なども交えながら、世界経済フォーラム（WEF）のAI部門を統括するKay Firth-Butterfield氏と、日立製作所の鈴木教洋研究開発グループ長が語り合った。



Kay Firth-Butterfield

AIガバナンスの第一人者。法廷弁護士、裁判官、大学教授を経て、AIと法律に関するLord Chief Justice's Advisory Panelのメンバー、AI Globalの共同設立者となる。UNESCO International Research Centre on AI and AI4Allのアドバイザーボードも務める。法律および国際関係の上級学位を有し、AIの活用によって生じる技術的、経済的、社会的な変化のさまざまな側面について発信を続ける。また、「リーダーシップのあり方を変える10人の女性」の一人として、ニューヨークタイムズで紹介されている。

期待が高まる AIによるブレークスルー

――近年、COVID-19や気候変動などの社会課題や世界情勢における不確実性を背景に、AI（Artificial Intelligence：人工知能）の活用が社会や暮らしの中に浸透してきています。今後、AIの活用がますます広がることによって、どんなことが期待できますでしょうか。

Firth-Butterfield 責任を持って正しく活用すれば、AIを用いて世界的な課題に取り組むことは十分に可能だと思います。しかし実際には、AIについて心配する人も増えています。欧州議会の調査によると、欧州では87%の人がAIを「非常に恐れている」という結果が出ています。米国で行われた同様の調査でも78%に上ります。ガバナンスに関して不確実性があるがゆえに、現時点でAIは従来予想されていたとおり普及しているとは言えません。ですが、そもそもAIは魔法の杖ではありませんから、AIですべての問題が解決できると考えて、それ以外の解決策を探すのをやめるべきではないと考えています。

鈴木 日立の社会イノベーション事業は、社会価値、環境価値、経済価値の三つの価値を向上させ、人々のQoL (Quality of Life) とお客様の企業価値を高めるSociety 5.0の実現をめざしています。日立は、デジタル技術を用いて、OT領域とIT領域の両方で、お客様やパートナーとのオープンイノベーションを通じて、社会や産業をよりインテリジェントなシステムに変えていきます。近年、人間の意思決定のサポートにAIを活用することで、非常に大量のデータに基づいた意思決定ができるようになりました。また、AIの活用によりビジネスや社会を変革することも可能になっています。そのため、イノベーションの源泉となるAIの重要性は近年ますます高まっています。

AIは日立のプラットフォームであるLumadaのコア技術であり、社会の課題を解決するための協創に欠かせない技術なのです。AIの産業応用は、鉄道分野での運行管理、エネルギー分野での送電・配電計画の最適化、医療・製造分野での機器メンテナンスの高度化、金融分野での融資審査など、幅広い分野に広がっています。実際に、AIを使って課題を解決するプロジェクトは数百件を超えています。環境問題や社会のレジリエンス、価値観の多様化に伴うQoLの向上など、社会が取り組むべき課題は、より高度で複雑化しています。こうした課題の解決に際して、AIは大きな役割を果たすでしょう。今後、AIは社会の幅広い分野で活用されることとなりますが、特に日立は「環境」、「レジリエンス」、「安心・安全」を事業領域と捉えており、これらの領域においてもAIが事業や社会の変革を牽引していくと考えています。

AIがもたらすリスクと、 新たな倫理の確立

—AIへの期待が高まる一方、AIの負の側面もクローズアップされています。それについてどのようにお考えでしょうか。

Firth-Butterfield AIの倫理原則を最初に打ち出したのは日本だったと思いますが、現在、世界中の国や企業がそれに倣っており、中でも約10の項目はすべての国で共通しています。すなわちこの共通項目は、AIの責任ある展開、有益かつ公正な活用という観点で極めて重要な核となるものなのだと思います。アルゴリズムを開発する際には、ダイバーシティやインクルージョンを考慮する必要がありますし、堅牢性や安全性、セキュリティも重要です。

AIは今後、より汎用的な存在に変わっていくでしょう。スピーディーに方針を策定し、事業活動を遂行するために

は、すべての企業はAI企業に変わっていかねばなりません。

近年、私たち世界経済フォーラム (WEF: World Economic Forum) がAIを活用している分野の一つに高齢者の支援がありますが、当事者である高齢者にとっては、それが安全で、かつ人間の尊厳を守るものであるかどうか非常に重要です。また、同じく社会的に弱い立場にある子どもたちのためのAI活用のガバナンスを考える「Smart Toy Awards」にも取り組んでいます。私たちは、AIのもたらすメリットを活用しながら、問題点を解決していかなくてはなりません。

鈴木 日立はまさにAI企業をめざしています。日立の社会イノベーション事業には、社会インフラや公共サービスに関わるものが多くあり、AIの異常行動や外部からの悪意ある行為は、社会全体や人命に重大な影響を与える可能性があります。また、AIが誤って発展すれば、差別や偏見、格差を助長する可能性もあります。ですから、公平で信頼できる社会を実現するためには、AIがどのように振る舞うのか、どのようなリスクがあるのかを正しく理解することが必要です。



鈴木 教洋

1986年東京大学大学院工学系研究科修士課程修了、日立製作所入社。デジタル画像信号処理、組込みシステムなどの研究開発に従事後、2012年日立アメリカ社シニアヴァイスプレジデント兼CTO、2014年中央研究所所長、2015年研究開発グループ社会イノベーション協創統括本部長を経て、2016年から現職。工学博士。映像情報メディア学会会員、電子情報通信学会会員、IEEE Senior Member。

日立は、歴史的にも社会インフラや公共サービスとの関わりが深い企業です。AIが普及し始める前から、技術が社会に与える影響に高い感度を持ち、技術者向けに倫理観の教育を行ってきました。しかし、AIや機械学習などの新しいテクノロジーにおいては、その特性を理解し、適応していくことが必要です。AIの進化と応用のスピードは非常に速く、これに伴って社会の変化も加速します。例えば、金融における自動売買や、SNSの世論への影響など、実際に社会のあり方が大きく変わり始めています。社会が間違った方向に進まないよう、AIの開発と応用を考えていくことが重要です。

——AIのリスクを低減し、安心・安全を担保した形で活用するためには、どのようなアプローチが必要でしょうか。

Firth-Butterfield WEFでは、マルチステークホルダーのアプローチを採用しています。AI活用によって生じるメリットとデメリットについて、正確な認識と、安心・安全な利用方法を普及させるため、各国政府の理解を促すことも私たちの課題の一つです。

企業において、AIを採用することは企業全体の問題です。AIをうまく活用するためには、組織体制を見直し、実際にAIを適用する製品についても検討しなければなりません。AIの研究にあたっては、いかにして多様性に富んだチームをつくり、製品がどのように使われるのかを考えることが課題です。これに加えて、社内教育を通じて、AIが自分たちの仕事を奪うものではなく、助けるものであることを社員に理解してもらう必要もあるでしょう。

非営利団体では、教育や医療の分野でAIを活用するために、財団が多額の資金を投入するケースが増えています。医師の数が、国民2万7,000人に対して医師1人の割合であるインドでは、チャットボットを活用して医療や治療の優先度などに関する情報を提供する活動を行っています。これに際しては、医療倫理の専門家とAI倫理の専門家を集めて独自のフレームワークを作成し、インドと同様に医師が不足している他の国・地域でも活用されています。

鈴木 政府の研究機関などがさまざまなガイドラインを公表しています。日立はそれらのガイドラインを参考にしながら、さまざまなアプローチで、AI倫理・AIガバナンスの取り組みを行っています。まず、社会におけるステークホルダーとの対話や合意形成を重視しています。また、AIを正しく開発・活用するための人財育成に力を入れています。そして、AI関連の研究や事業活用のリスクを分析・評価・管理するためのリスクマネジメントの仕組みや、AIを活用

したサービス・プロダクトの品質保証の仕組みを整えています。さらに、「優れた自主技術・製品の開発を通じて社会に貢献する」という日立の企業理念に基づいて、AI倫理を担保するためのAI技術およびデータ処理・管理技術の研究開発などが重要と考え、実践しています。

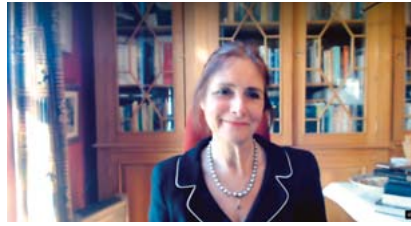
これらを実践するためには、すべての基礎となるAIの開発・活用に関する企業方針を定め、全社的な意識の徹底を図る必要があります。そこで日立は、社会イノベーション事業ならではの「AI倫理原則」を独自に策定し、公開しています。その特徴は、社会・産業インフラへのAI活用を視野に入れて、「計画」、「社会実装」、「維持管理」の3段階の行動基準を定めている点にあります。

人財育成の面では、Lumadaによるデジタルイノベーションを加速させるべくデータサイエンスのトップ人財を集めたLumada Data Science Lab.を設立し、AIやデータサイエンスの知識とスキルを磨いています。現在、データサイエンティストなど約3,000人の専門家を育成しているほか、幅広い職種や部門を対象に、AI倫理に関する社内教育を討論会や座学の形で実施しています。リスクマネジメントについては、日立のAI倫理原則に対応する形で、リスクを評価するためのチェックリストを策定しています。チェックリストを用いたリスク評価は、協創を通じた研究の開始時や、PoC (Proof of Concept) の各フェーズでも実施しています。また、外部の有識者で構成されるAI倫理アドバイザリーボードを設置し、社外からの意見を取り入れるようにしています。また、AI固有の特徴を踏まえた開発・品質保証のガイドラインも策定しています。

新たなテクノロジーに「信頼」を埋め込むガバナンス

——新しいテクノロジーを導入するためには、ガバナンスを強化する必要があるというのが世界の共通認識になりつつあります。そのためには、どのような取り組みが必要でしょうか。

Firth-Butterfield この問題への取り組み方は、国や地域によって異なります。例えば、欧州では、リスクベースのAIに対する規制法案であるAI Actが提案されました。AI Actでは、AIを活用した顔認証はリスクが高く、ガスをより適切に利用するためのAI活用はリスクが低いというように、リスクに応じたAI活用に向けた規制が提案されています。ただ、AIに関する法制化については懸念もあります。さまざまな分野でAIの導入が進む一方、各国政府が法律を制定するには長い時間がかかってしまいます。これに関



しては、カテゴリーを設けて認証を検討するという欧州の方式は、賢明だと思います。逆に法律などを設けず、自由放任主義的なアプローチを採っているケースもありますが、ガバナンスや法律がイノベーションの妨げになると言われがちな米国においても、連邦政府内でAI関連の法律や調達について考える人々が現れつつあります。上院の党員集会では、AIの調達について議論が始まっていますし、国防総省でも、AIの調達戦略について話し合いが進んでいます。また、機会均等委員会でも、人財管理におけるAIの導入を慎重に検討しています。

鈴木 私は、テクノロジーとガバナンスは車の両輪であると考えています。ガバナンスシステムが整っていれば、テクノロジーへの信頼と安心感が生まれます。設計段階からしっかりとガバナンスを考慮することで、技術やイノベーションの社会への適用を加速させることができるのです。

テクノロジーのガバナンスについては、プライバシーや安全性、セキュリティだけでなく、透明性や包括性、アカウントビリティについても考えていく必要があります。また、機械学習などのAI技術は、データから学習するため、データの偏りによる公平性の欠如がないように注意する必要があります。そこで、日立はAIの倫理的利用を導く原則として、七つのポイントを挙げています。(1) 安全重視、(2) プライバシー保護、(3) 公平性重視、(4) 適正開発・利活用、(5) 透明性・説明責任重視、(6) セキュリティ重視、(7) 法令遵守です。テクノロジーガバナンスを構築するためには、複数のステークホルダーからのさまざまな視点が必要であり、またガバナンスを維持するためには、常に見直すことができる柔軟性と機動性が重要となります。そのような条件を企業のフレームワークに組み込むことがカギになるでしょう。

Firth-Butterfield ガバナンスに関しては、ESG (Environment, Social, Governance) の仕組みの中にAIの影響を含めるということに、私たちも取り組み始めています。投資家やVC (Venture Capital) 企業からは「倫理的

なAI企業やスタートアップに投資するために、何を基準に判断すればいいのか」という声が寄せられています。ガバナンスという大きな枠組みの中で、私たちはそれぞれの方法で考えていかなければならないと思いますが、特に上場企業にとっては、社会からの要請としてのAIに関する倫理について強く意識しなければならなくなるでしょう。

——社会を円滑に機能させるための仕組みとして、「信頼(トラスト)」が再び注目を集めています。テクノロジーにおける信頼についてはどう思われますか。

鈴木 2021年4月に行われたWEFのグローバル・テクノロジー・ガバナンス・サミットにパネリストとして登壇した際、テクノロジーに対する信頼というテーマについてお話ししました。第四次産業革命、特にデジタル技術の導入は、社会を大きく変えましたが、世界情勢における不確実性の高まりによってこれまで信頼を築いてきた方法が機能不全に陥り、信頼を築くための新たな方法を模索することになりました。近年のDX (デジタルトランスフォーメーション) は、過去に比べて急速に進んでいることが分かっています。そのため、技術の信頼性は、人権や環境に関連したSDGs (Sustainable Development Goals) の達成にも関係してきます。日立にとって、技術の信頼性を確保することは、事業の根幹に関わるものですので、WEFと経済産業省の協力を得て、トラスト・ガバナンス・フレームワークに関するホワイトペーパーを発表しました。これは、信頼を築くための新しいガバナンスのあり方を議論するための基礎となるものです。このホワイトペーパーでは、新たな信頼の構築にはその技術が信頼に値することを示す客観的事実や証拠を蓄積し、誰もがそれを確認できるようにし、社会の認知度を高めることが必要であると提案しています。

このように日立は、AIを正しく開発・利用するだけでなく、指針を公開し、その倫理指針に基づいてAIが実装されている事実を開示することで、日立のAIに対する信頼感を高めようとしています。

Firth-Butterfield 今日、世界中のどの国においても、テクノロジーに対する信頼は大きな課題です。テクノロジーを信頼しているがゆえに、人々の分断が進んでいることを目の当たりにしています。だからこそ、人々には実際に信頼できるテクノロジーを信頼してもらいたいです。ソーシャルメディアで語られる情報の中には、多くの間違いやフェイクニュースがありますし、専門的なテクノロジーに関して、一般市民の人々に説明し、正しく理解してもらうことは容易ではありません。

新たな信頼について特に悩ましいのは、人々が他の人間よりもコンピュータや機械を信頼してしまいがちだということです。AIをはじめ、新たなテクノロジーによって、これまで人類が築き上げてきた常識が覆されるという可能性について、私たちは十分な議論を重ねてきたとは言えないでしょう。保釈の審問で裁判官がコンピュータが出した答えを信じて判決を下しても、銀行がアルゴリズムに基づいて融資の可否を決定しても、そのアルゴリズムが信頼できる方法で設計されていれば問題は生じないはずですが、そのためには倫理が何より重要です。AIを設計する企業が、倫理的な観点からAIを設計し、私たちが使用するものに信頼を埋め込むことは、非常に大きな意味を持つと思います。

カギとなるのは、 ステークホルダーとの協創

——AIの倫理やテクノロジーのガバナンスは、一つの組織だけに収まるものではありません。さまざまなステークホルダーの中でのコラボレーションが重要と思われませんが、いかがでしょうか。

鈴木 テクノロジーの進化のスピードは非常に速く、従来のように政府がルールを定め、私たちが従うというガバナンスのあり方では不十分になっています。ですから、さまざまなステークホルダーが知恵を絞り、ガバナンスが機能しているかどうかを常にチェックし、アップデートしていくことが欠かせません。この中で、われわれのような技術系企業が重要な役割を果たすことは間違いありませんが、企業が単独で信頼を築くことは不可能です。ですから、政府機関、学術的・科学的に中立な立場にある第三者機関、専門家ではない市民、そしてその集合体であるコミュニティがガバナンスに参加することが重要になってきます。

日立は、今後も、定例会議やワークショップなどを通じてWEFの活動に参加し、国際協力を通じて複数のステークホルダーとの連携を図っていく考えです。

Firth-Butterfield 私たちも同じ考えです。私が2017年にAIチームを立ち上げるためにWEFに参画した理由の一つは、WEFでは、世界をリードするマルチステークホルダーのコラボレーションが可能だからなのです。私たちは、日立のようなビジネスパートナーに恵まれており、政府や学術機関、そして市民やコミュニティと継続的な関係を築くことができている。これは、まだ道半ばにある世界中の企業をサポートし、信頼を築くうえでとても重要なことでしょう。

一方、責任あるテクノロジーの利用方法を確立するためには、AI倫理担当責任者といった枠組みを構築することが重要になります。企業がAI企業へと発展していく過程では、この課題を重く受け止め、検討している経営者も少なくはないでしょう。日立の経営役員には、技術者、研究者の方も多いと思います。今日語り合った倫理や信頼といった事柄は、民間企業の人々と共有したい内容でもあります。WEFではAIなどの最新デジタル技術を駆使したスマート工場「Lighthouse」のプロジェクトも手掛けていますが、その中でも倫理的なAIは重要なテーマになっていくと思います。

鈴木 責任あるAIが強く求められていることを、私も痛感しています。日立は、社会イノベーション事業に取り組む企業として、人財、システム、テクノロジーという面から、責任あるAIの実現をリードし、よりよい未来社会の構築に貢献していきたいと考えています。めざすのは、誰もがAIを信頼し、AIにさまざまな社会的役割・機能を任せることができる世界です。そうしたAIが正しく開発・活用されれば、社会の透明性や公平性を高めることにもつながっていくでしょう。社会でどのようなAIが求められるかはもちろん用途や領域に応じて異なります。AIに大きく依存する場合もあれば、AIだけに任せるのではなく、人間も含めたシステムとなる場合もあります。ですから、人間とAIが共存するというビジョンを持ち続けることは、AIの価値を高めていくうえで不可欠なのだと考えます。本日はありがとうございました。