

社会イノベーション事業のための 日立のAI倫理原則とその実践

日立の広範な事業領域でAIを活用していくにあたり、その社会に与える影響は重大であるため、AIのガバナンスおよび倫理は、日立にとって重要なアジェンダである。従来、安全、プライバシー、公平性、透明性・説明責任、セキュリティなど、さまざまなAI倫理上の観点が指摘されてきた。欧州のAI規制案や経済産業省のAIガバナンスガイドラインが公開されるなど、原則のみならず、ガバナンスの実践へ向けた動きが急速に活発化している。

日立は、2021年2月に「社会イノベーション事業にAIを活用するためのAI倫理原則」を発表した。またAI倫理原則に基づき、研究・PoC・製品開発・保守運用に至る各段階で、ガバナンス確立のための取り組みを実践している。

本稿では、日立のAI倫理原則の根底にある考え方について述べ、その実践について概説する。

吉田 順 | Yoshida Jun

柳井 孝介 | Yanai Kohsuke

1. はじめに

イノベーションの源泉としてAI (Artificial Intelligence: 人工知能) の重要性は拡大しており、鉄道、エネルギー、医療、製造、金融など幅広い分野へと適用されつつある。環境問題、社会のレジリエンス性の向上、多様化する価値観の中でのQoL (Quality of Life) 向上など、社会が取り組むべき課題は高度化し、複雑化している。これらの社会課題の解決においてAIの果たす役割は大きいと考えられ、将来、社会の広範な領域にAIが活用されていく。

日立はLumadaを立ち上げ、社会イノベーション事業におけるデジタル技術の活用を加速している。Lumadaにおいても、新たな価値を生み出す原動力としてAIが欠かせない存在になっている。日立の社会イノベーション事業は、社会インフラや公共性の高いサービスに関わるものが多

く、ここにAIを活用していくときの社会への影響は極めて大きい。公平で信頼できる社会の実現に向けては、AIの振る舞いやそれが内包するリスクを正しく理解し、正しく適用・実装・運用していくことが必須となる。

日立では、AIの開発や社会実装に向けたコーポレートポリシーとして、2021年2月、「AI倫理原則」を策定した¹⁾。同時にAI倫理に関する取り組みをまとめたホワイトペーパーも公開している。社会イノベーション事業を担っている日立ならではの行動基準や実践項目を定めた。

日立は創業期から人間中心の視点を受け継いできている。例えば、日立製作所誕生の母胎となった日立鉱山では、煙害対策のために1915年、製錬所に当時高さ世界第1位となる約156 mの大煙突を完成させた。またその周囲の山々に観測所を設け、大気の状態と精錬生産量を一部連動させた社会に開かれた経営システムを実現した。また2000年には、医療機器企業としては初めて倫理審査委員会を設立し、医学系研究や脳科学研究において、いち早く倫理に取り組

んできた。さらに2014年7月には、プライバシー保護対策を統括する「パーソナルデータ責任者」,「プライバシー保護諮問委員会」を設置し、プライバシーの保護にも配慮したデータ利活用に取り組んできた。

「優れた自主技術・製品の開発を通じて社会に貢献する」ことを企業理念としてきた日立にとって、倫理は企業文化として深く根付いているものであり、AIのみならず、日立の技術・製品開発において不変の前提としてあったものである。日立のAI倫理原則は、それを改めてAIの文脈で明文化したものであり、社員の意識の統一化と社内ガバナンス体制構築の指針となるものである。

以下、2章では、日立のAI倫理原則とその根底にある考え方を説明する。3章では、AI倫理原則の実践について述べる。

2. 日立のAI倫理原則

日本においては、2019年3月に「人間中心のAI社会原則」が公表されている。国際的には、2019年5月にOECD（経済協力開発機構）においてAI原則を含む理事会勧告が採択され、また2019年6月にG20 AI原則が採択されている。各国、各社の取り組みは、AIネットワーク社会推進会議「報告書2021」にまとめられている²⁾。AIの倫理の問題はグローバルで認識され、国際的な議論を経て、多様な観点を取り

込みながら論点が整理されてきている。

日立は、これらの国際的な議論を踏まえつつ、日立の企業文化と社会イノベーションという事業の性質に合致した独自のAI倫理原則を策定している。日立のAI倫理原則では、「計画」、「社会実装」、「維持管理」の三つのフェーズにおける行動規準と、すべてのフェーズに共通する七つの実践項目を定めている（図1参照）。

AI倫理原則の策定にあたっては、10以上の事業部と議論を重ね、日立の文化・事業に合っているかを確認した。また日本・米国・アジア・欧州の研究部門にて、事前にAI倫理原則を反映したリスクアセスメントを試行して、AI倫理原則に実効性があるかを確認した。このようなプロセスを経て、地域や事業部をまたがって、共通の理念を表す共通言語としてAI倫理原則を策定した。

またAI倫理原則にひもづける形で、リスクアセスメント用のチェックリストや教育教材を整備し、AIガバナンスのための運営を行っている（図1参照）。本章では、AI倫理原則の行動基準と実践項目について説明する。

2.1

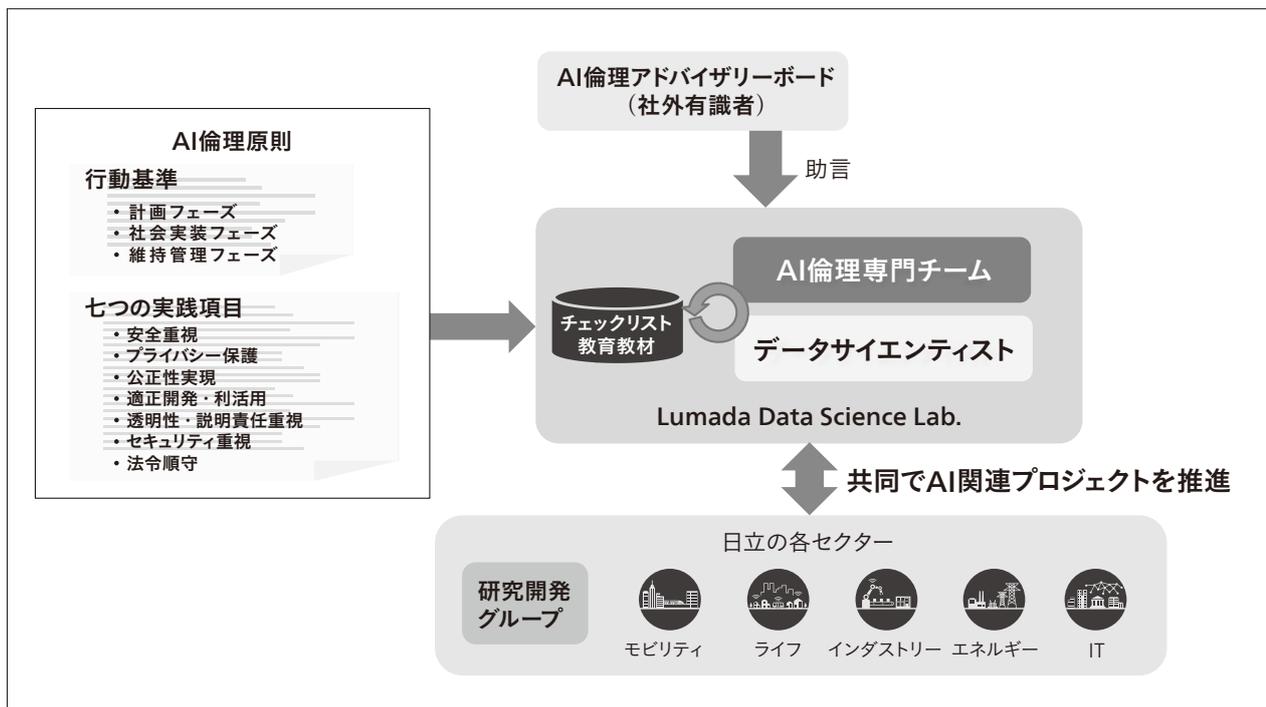
三つのフェーズにおける行動基準

(1) 計画フェーズ：「持続可能社会実現のために、AIの開発、利活用を計画します」

社会イノベーション事業においては、制度や法律などの社会システムにも注目して、社会や経済の新しい関係性を

図1 | 日立のAI倫理に対する取り組み概要

日立のAI倫理原則では、「計画」、「社会実装」、「維持管理」の三つのフェーズにおける行動規準と、すべてのフェーズに共通する七つの実践項目を定めている。



つくっていく³⁾。このような社会的挑戦においては、最初にアイデアが出て終わりではなく、ある程度の時間軸の中で、社会での発見的・探索的な試行錯誤のプロセスを経て構想をつくり出していく。このある程度の時間幅を持った構想過程において、倫理的な観点で、適切な目的のためのAI活用を計画することが必須である。

そのため日立は、サービス・ソリューションの提供時のみならず、社会イノベーション事業の計画フェーズにおけるAI倫理観点での検討を重視する。研究開発、PoC(Proof of Concept)を含むイノベーション創生の初期のフェーズから、AI倫理を実践していく。

(2) 社会実装フェーズ：「人間中心の視点で、AIを社会実装します」

社会イノベーション事業におけるAI活用は、社会へのインパクトが極めて大きいことを認識し、責任をもって社会実装を進めることが重要である。これは日立が創業以来、受け継いできた精神である。

従来技術と比較して、AIは、(a) 自動化・無人化を押し進めること、(b) 振る舞いを人が明示的な形で決めるのではなく、学習データに依存して振る舞いが決まること、(c) 統計的な振る舞いをするなど、などが特徴である。これらの特徴を深く理解し、AIのアルゴリズムやAIを活用するシステム側の品質保証はもとより、社会への影響の大きいユースケースにおいては、学習用データの健全性やAIおよびシステムの透明性などにも考慮して実装することが必須である。このときに起点となるのが人間中心の視点であり、社会の人々に必要とされているものは何なのかを、常に第一において実装を進めていく。

(3) 維持管理フェーズ：「提供価値が長期間にわたり持続するよう、AIを維持管理します」

社会インフラには、数十年のタイムスパンで使われるものもある。そのため日立の事業においては、AIを社会実装した後も、比較的長い期間でAIの影響を考えていくことが必須である。

例えば、環境変化などが原因で、時間が経過するとAIへ入力されるデータに変化が生じ、それがシステム全体に影響を及ぼすことが考えられる。そこで、AIを含むシステムが長期かつ安定的に価値を提供できるよう、AIを維持管理することが必要である。

これは、単にAIが異常動作しないようにAIを監視するというに留まらない。社会への提供価値という観点を重視し、AIを含むシステムが引き続き社会にとって求められているものであるのかどうかを、モニタリングするよう努める。

2.2

七つの実践項目

「計画」、「社会実装」、「維持管理」の三つのフェーズに共通する実践項目として、「安全重視」、「プライバシー保護」、「公平性実現」、「適正開発・利活用」、「透明性・説明責任重視」、「セキュリティ重視」、「法令順守」の7項目を定めている。これらは、それぞれ具体的な確認項目へブレイクダウンされる。イノベーション構想、研究開発、PoC、ソリューション・システム開発など、さまざまなフェーズにおいて、この七つの実践項目が参照され、AI倫理原則にのっとった企業活動になっているかをチェックできるようにしている。またこの七つの項目に基づいて、リスクの高いケースを収集・分類しており、教育やリスク低減策の策定などに役立てている。以上のように、この7項目が、日立のAI倫理実践における軸となっている。

3. 実践

AI・アナリティクス分野の中核組織であるLumada Data Science Lab. (以下、「LDSL」と記す。)の中に、日立におけるAI倫理のCoE (Center of Excellence)として、データサイエンティストやプライバシー保護の専門家などから成るAI倫理専門チームを設置している。このAI倫理専門チームが、AI倫理原則を遵守するための仕組みを構築し、AI倫理に関わるガバナンスを継続して改善する役割を担っている。またLDSLが中心となり、AI倫理を担保するための技術の研究開発や人材教育を推進している。LDSLには、日立内の社会イノベーション事業の先進の協創プロジェクトが多く集まるため、LDSL自体がAI倫理を実践する場となっている(図1参照)。

さらに、AI倫理の取り組みに対し、客観的に評価し、社外からの意見を取り入れながら改善するため、社外有識者から成るAI倫理アドバイザリーボードを設置している。アドバイザリーボードからは、AI技術視点のみならず、消費者視点や社会科学視点からも助言を得られるようにしている。

3.1

リスク管理

安全重視、公平性実現、透明性・説明責任重視などのAI倫理原則で定めている実践項目に対応したチェックリストを作成し、これを用いて個別案件ごとにリスクアセスメントを実施している。これにより、AIに関連する研究開発や、AIを活用する業務のリスクを効率的に抽出できるようにしている。チェックリストでアセスメントした結果に

基づき、推進上の留意点を助言したり、対応策の検討を支援したりしている。またチェックリストで明らかになった留意事項を踏まえ、マネージャーがその案件を継続するかどうかを判断できるようにしている。また必要に応じて、社外有識者から成るアドバイザーボードにて、個別の審議・助言を得るようにしている。

このリスクアセスメントを実施することで、広範に実施されているAIの研究開発やAIの利活用案件の情報を集約でき、ガバナンスのベースとなっている。またリスクアセスメントのプロセスを実施することで、AI倫理に関するナレッジを集約・蓄積していけるようになっている。

チェックリストを用いたリスクアセスメントはさまざまなフェーズで行われる。例えば、技術研究のフェーズに関しては、研究開始段階において、チェックリストを用いて研究テーマのリスクや妥当性を検証するスクリーニングをかけている。この研究フェーズに関しては、年間600を超える研究テーマに関してチェックリストを用いてスクリーニングを行っている。また事業適用段階においては、顧客案件に引き受け時から開発に至る各フェーズにて、チェックリストを用いたリスク管理を行っている。年間約100の案件に関してチェックリストを用いた審議を行っている。

3.2

AIを活用したサービス・システム・製品の品質保証

日立グループ全体でAIを活用したサービス・システム・製品の開発が増えていることから、AI固有の品質保証を強化している。産業技術総合研究所の機械学習品質マネジメントガイドライン（AIQM）⁴⁾、およびAI倫理原則に基づく、AI固有の開発・品質保証ガイドラインを策定し、日立グループ内に公開している。

これは、AI固有の品質保証に関する基本的な考え方、PoCおよび開発フェーズにおけるAIモデルの開発方法、運用フェーズにおける予測精度の監視、精度劣化時の再学習方法、AI倫理に関わる安全重視、プライバシー保護、公平性実現などを体系的にまとめたものである。開発者がガイドラインを参照して、AIを活用したシステムや製品の開発を行うとともに、従来の品質保証プロセスの仕掛けの中にAI固有の観点を付加したレビューを実施している。

3.3

AI人材の育成とAI倫理教育

AI倫理原則を遵守するにあたり、AI技術に対する深い理解が不可欠である。日立グループ全体で、データサイエンティストが集まるデータサイエンス部会を立ち上げ、データサイエンティストとして各自が持っている知見やノウハウ

を共有している。このデータサイエンス部会を中心に、AIやデータサイエンスの専門知識を持つデータサイエンティスト約3,000人を育成している。このように、AIの深い知識を備えた人材を多く育成することが、AIを倫理的に利活用していくためのベースとなる。

以上のAI人材の育成の上に、AI倫理の研修や社内の普及啓発活動を実施している。具体的には、上述のデータサイエンス部会においてAI倫理の勉強会を実施している。また、AI倫理の世界情勢、研究、社内ガバナンス、規制などに関して、横断的に最新動向を共有して議論する討論会を開催しており、職種や部門を問わず日立グループの社員であれば参加でき、経営幹部含む500名以上が参加している。座学での研修や勉強会とは別に、職場ごとでのユースケース討論型の教育も実施している。AIが活用されるユースケースを10ほど紙に記載して職場に配布し、各職場ごと10～20名程度の単位で、配布された資料の中からいくつかのユースケースを選択し、その具体的なユースケースにおけるAI倫理観点での問題点を議論する。AIの研究や活用に携わる1,000名以上が、本教育を受けている。

3.4

AI倫理を担保する技術の研究および開発

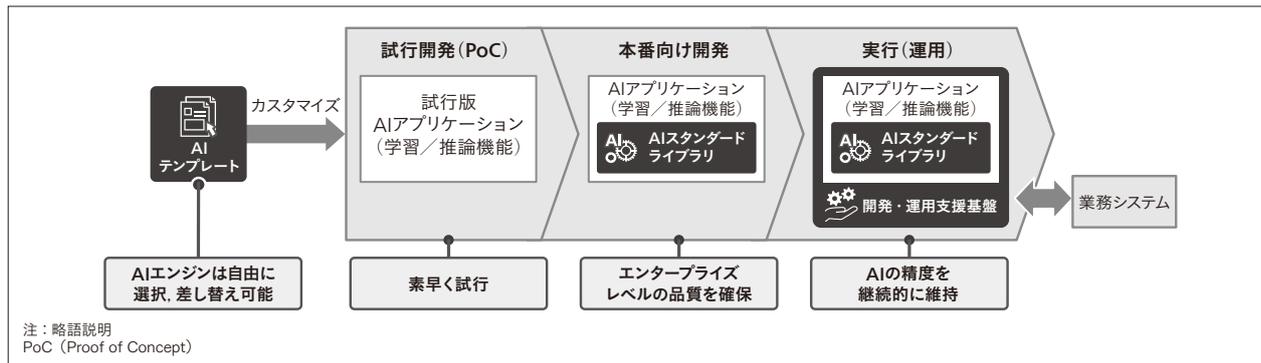
社内の仕組みのみならず、「優れた自主技術・製品の開発を通じて社会に貢献する」という日立の企業理念に基づき、AI倫理を担保するための技術の研究開発も推進している。説明可能なAI、データバイアスを検知・低減し公平性を高めるAIなどのAIのアルゴリズムに関わるもののみならず、データのトレーサビリティを支援するデータリネージなどのデータガバナンス技術、モデルのモニタリング・劣化検知技術などの研究開発も行っている。これらの一部は、別稿「AIのトラストとガバナンスを支える研究開発」⁵⁾でも紹介する。

また、AI倫理原則に基づき、企業や社会インフラへのAI実装を迅速化する「AIアプリケーションフレームワーク」を提供している⁶⁾（図2参照）。

本フレームワークは、AIに関する日立の開発・運用ノウハウを結集したものであり、高度な知見と専門性が必要とされるエンタープライズ領域において、高効率かつ高品質なAIシステムの開発・運用を支援する。具体的には、システム稼働後に、AIシステムへの入力データや結果を常時監視・分析し、予め設定したルールに基づいて異常なデータや結果を自動検知できる。これによりAIの予測精度の劣化を防止する。また精度が劣化した場合、新たに再学習を行ったモデルを呼び出したり、過去のモデルに切り替えたりすることが可能である。

図2 AIアプリケーションフレームワーク

本フレームワークは、AIに関する日立の開発・運用ノウハウを結集したものであり、高度な知見と専門性が必要とされるエンタープライズ領域において、高効率かつ高品質なAIシステムの開発・運用を支援する。



4. おわりに

本稿では、日立の社会イノベーション事業のためのAI倫理原則とその実践について述べた。

日立の文化・事業に基づいた独自のAI倫理原則を策定することで、全社としての取り組み姿勢の明文化、社員の意識の統一化が図れた。また社内でのガバナンス体制構築の取り組みも加速した。原則策定により社員のAI活用に関する意識が以前に比べて明らかに高まったと感じている。

AIガバナンスとイノベーションが両天秤にかけられ、AIガバナンスを強化するとイノベーションが阻害されるという意見もある。しかし、イノベーションのためにこそ、AI活用のための原則やガバナンスが必要で、国際協調・ステークホルダー間での協調の下に、社会の合意のうえでAIの研究開発・活用を進める必要がある。日立は過去にも、CMOSイメージング計算機^{7), 8)}やディベート型AI⁹⁾などのAI技術に関して、ステークホルダーダイアログを実施した。また2021年にもAIガバナンスやトラストをテーマとして、「協創の森ウェビナー」¹⁰⁾を実施し、社外の視点を研究開発に取り入れようと試みている (図3参照)。

今後、マルチステークホルダー間での協調をより一層意

図3 協創の森ウェビナーの様子

2021年10月、社内外の有識者とAIのガバナンスについて議論するウェビナーを、東京・国分寺の「協創の森」より発信した。



識して進めることで、社会価値・環境価値・経済価値を向上するとともに、人間の尊厳に立脚した快適で持続可能な社会の実現と、世界の人々のQoLの向上に貢献していく。

参考文献など

- 1) Lumada Data Science Lab., AI倫理, <https://www.hitachi.co.jp/products/it/lumada/about/ai/ldsl/index.html#p05>
- 2) 総務省, AIネットワーク社会推進会議 報告書2021 ~「安心・安全で信頼性のあるAIの社会実装」の推進~, https://www.soumu.go.jp/main_content/000761967.pdf
- 3) 紺野登, 野中郁次郎: 構想力の方法論, 日経BP (2018.7)
- 4) 産業技術総合研究所, 機械学習品質マネジメントガイドライン 第2版, (2021.7)
- 5) 内田尚和, 外: AIのトラストとガバナンスを支える研究開発, 日立評論, 特別増刊号, 20~27 (2021.11)
- 6) Justware AIアプリケーションフレームワーク, <https://www.hitachi.co.jp/products/it/appsdiv/service/justware/ai-apfw/index.html>
- 7) 山岡雅直, 外: 社会システム最適化に資するCMOSイメージング計算機, 日立評論, 99, 3, 329~333 (2017.5)
- 8) 中村知倫, 外: AIアプリケーション開発と運用を支援するJustware AIアプリケーションフレームワーク, 日立評論, 103, 6, 649~653 (2021.11)
- 9) 柳井孝介, 外: ディベート人工知能 AIの基礎研究, 日立評論, 98, 4, 269~272 (2016.4)
- 10) 日立製作所 協創の森ウェビナー, https://www.hitachi.co.jp/rd/open/kyosonomori/collaboration/webinar_2021.html

執筆者紹介



吉田 順
日立製作所 Lumada Data Science Lab. 所属
現在, AI倫理専門チームの取りまとめ, AI・アナリティクスを活用した顧客協創, 日立社内のデータサイエンティスト育成に従事



柳井 孝介
日立製作所 研究開発グループ 先端AIイノベーションセンタ
メディア知能処理研究部 所属
現在, メディア知能処理の研究開発に従事
博士 (科学)
人工知能学会会員