

ACTIVITIES

社会から信頼されるAIを実現するためのAIガバナンス

経済産業省
商務情報政策局
情報政策企画調整官

泉 卓也

株式会社 日立コンサルティング
スマート社会基盤コンサルティング
第2本部 ディレクター

美馬 正司

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 部長

柳井 孝介

AI原則の実践に向けた動きが始まっている。経済産業省では、2021年7月に、AIの社会実装の促進に必要なAI原則の実践を支援する目的でAIガバナンス・ガイドラインを公表した。日立においても、AI倫理原則を策定し、その実践に向けた取り組みを開始している。

2021年10月に開催されたHitachi Social Innovation Forum 2021 JAPANでは、AIガバナンス・ガイドラインの作成をリードされた経済産業省の泉卓也氏を招き、日立グループのキーパーソン二人と、AI倫理を実践していくにあたり、政府や企業を含むマルチステークホルダー間でどのように協調すべきかについて意見を交わした。



【セッションⅠ】

講演（1）

リスクを管理し、正のインパクトを最大に
AIガバナンス・ガイドラインについて （泉 卓也）

イノベーションの源泉としてAI（Artificial Intelligence：人工知能）の重要性が拡大する中、経済産業省では、2021年7月9日にAIガバナンス・ガイドライン[※]を公表した。まず、この背景を説明したい。

AIは社会に多大なる便益をもたらしてくれる。例えば、画像認識のAIは自動運転の開発やX線画像の読影などに活用されている。その一方で、AIシステムをつくるための学習用データ自体に差別などのバイアスが反映されていて、AIシステムを通じて、それが露呈することもあり得る。AIのガバナンスに関する取り組みの歴史は、AIの有効活用を通じた便益の享受とネガティブな側面への対処を両立するための取り組みの歴史に他ならない。

2019年3月に、この取り組みを前に進めるために、ハイレベルな指針である人間中心のAI社会原則が内閣府より公表された。AI社会原則はプライバシー、セキュリティ、公平性、説明責任、透明性などの七つの原則から成る。2020

※）経済産業省 AI原則実践のためのガバナンス・ガイドライン ver. 1.0
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/2021070902_report.html

年6月には、このようなハイレベルの指針に沿った取り組みを実施しやすくするためのAIガバナンスの議論が政府内で本格的に開始した。経済産業省では、2021年1月に「我が国のAIガバナンスの在り方」をまとめて、パブリックコメントを募集している。そして、これらの意見を踏まえ、7月にAIガバナンス・ガイドラインを公表した。

経済産業省の検討では、「AIガバナンス」を、AIの活用によって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ、そこからもたらされる正のインパクトを最大化することを目的とした、ステークホルダーによる技術的、組織的、及び社会的システムの設計及び運用と定義している。ここでは社会的システムについて説明したい。

図1に示すように、AI原則で示される社会のめざすべき価値は図の「What」に位置づけられる。現在の課題は「How」の部分になる。実は、このAI社会原則を実現する方法はたくさんある。例えば、AIに特有の機能に着目した横断的なルールを用いることが考えられる。また技術標準を決めていくという手法もある。さらには、自動車や医療などの分野別にルールを決めていくことも考えられる。正のインパクトを最大化し、負のインパクトを最小化するために、これらをどのように組み合わせるかが、AIガバナンスの課題となる。経済産業省では、横断的なルールについて検討し、法的拘束力のないガイドラインを提供すべきという結論に至った。そうして策定されたのが、経済産業省のAIガバナンス・ガイドラインである。

AIガバナンス・ガイドラインは、企業レベルのAIガバナ

ンス実践の指針としてまとめたものである。AIガバナンス・ガイドラインは、同じく経済産業省で推進しているアジャイル・ガバナンスのフレームワークを活用している。技術の発展が途上であり、社会の受け止め方も変化しているAIのような分野のガバナンスの構築に有効なフレームワークである。

図2に示すように、AIガバナンスは二重ループから成る。まず、環境・リスク分析から始まる。AIの場合では、AI技術を知り、社会の受け止め方を知り、自己を知ることが重要になる。この分析を踏まえ、ゴールを設定する。それぞれの企業が大切にしていける価値を特定していく。その後、ゴールを達成するためにAIマネジメントシステムを構築する。AIの場合には、設定したゴールを達成するための、ゴールからの乖離を特定するための乖離評価が重要になる。また、リテラシーの向上や事業者間連携も重要である。ループの一番下に記載している運用においては、AIマネジメントシステムや個々のAIシステムの状況をモニタリングする。必要に応じて、外部に開示していくことも大切である。そして、ループの中央の評価においては、AIマネジメントシステムが適切に機能している否か、機能していないとすればどのように改善すべきかを検討する。事業環境が大きく変化する場合には、環境・リスク分析を再度行い、ゴール設定から見直す必要が出てくる。AIガバナンス・ガイドラインは、企業がステークホルダーと対話しながら継続的に改善できるように設計している。このガイドラインを伴走相手として、少しずつでも実践していただければ幸いである。

図1 AIガバナンスの全体像

AIガバナンスの社会的な側面は、(1) ゴール、(2) 横断的で中間的なルール、(3) 個別分野等にフォーカスしたルール、(4) モニタリング／エンフォースメントという層構造で考えると分かりやすい。

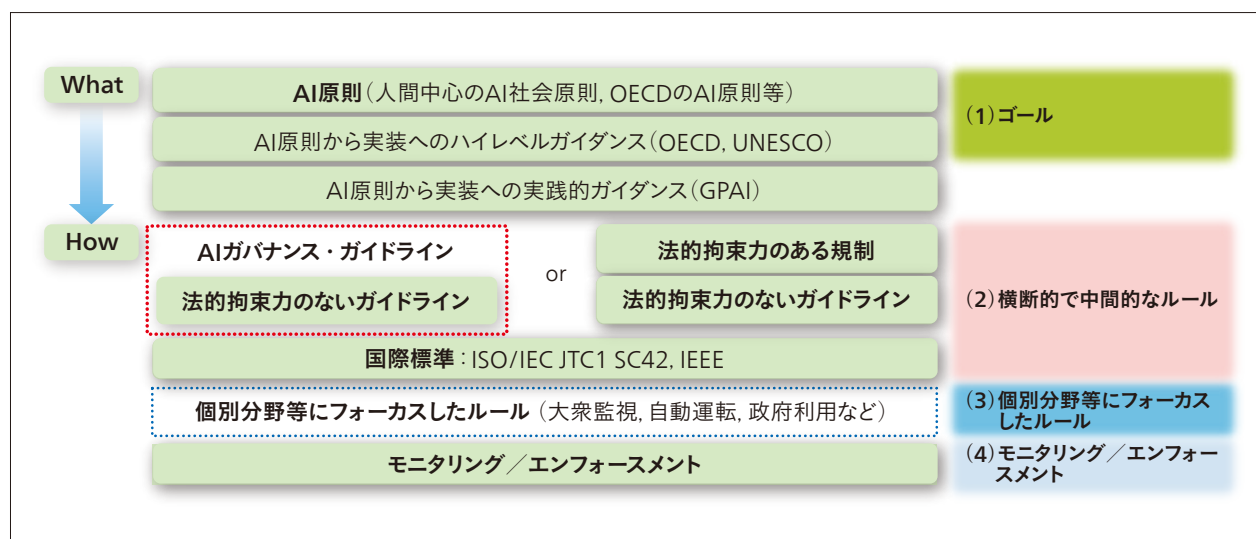
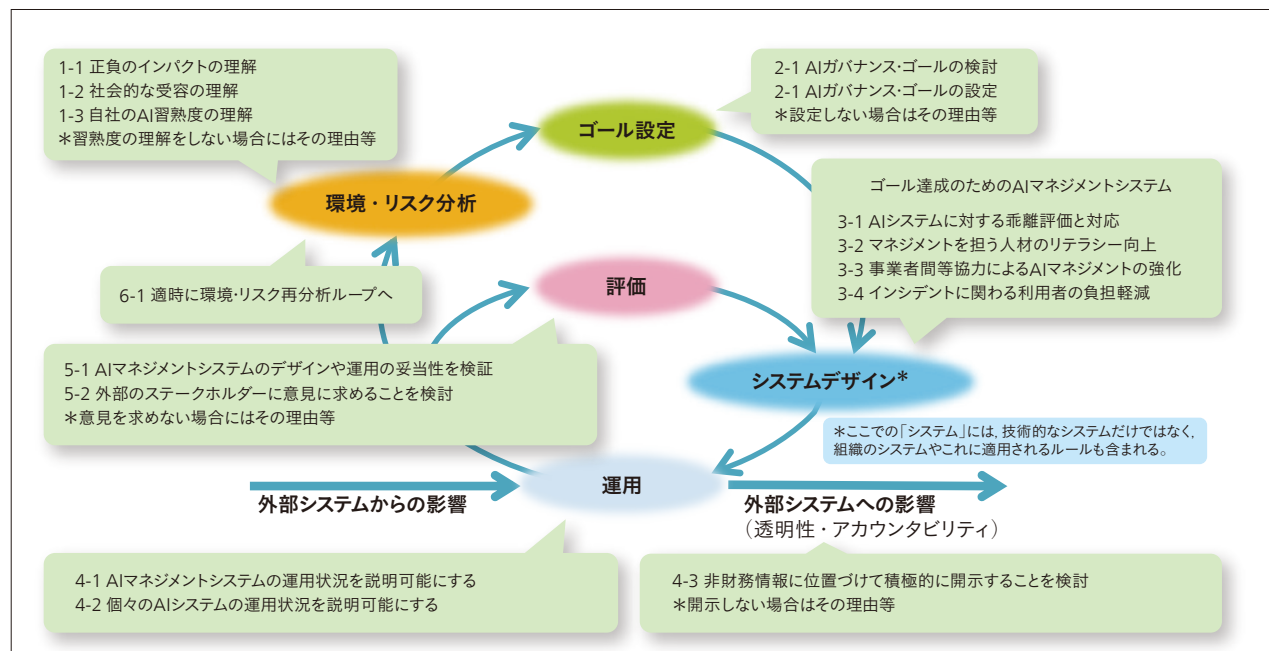


図2 AIガバナンス・ガイドラインの全体像

AIガバナンス・ガイドラインは、企業レベルのAIガバナンス実践の指針としてまとめたものである。



講演（2）

「AI倫理・バイ・デザイン」をめざして

日立におけるAIガバナンスの取り組み（美馬 正司）

日立においてもAIガバナンスの取り組みを開始している。AIに関しては変化のスピードが速い領域だと認識しており、いかに柔軟かつ迅速にガバナンスの仕組みを適応させていくかが重要だと考えている。経済産業省のアジャイル・ガバナンスのフレームワークに対応付けて、日立の取り組みを述べたい。

まずは、図2の二重ループにおけるゴールとして、日立では、AIに関する環境・リスク分析等の知見を積み重ねたうえで、AI倫理原則を策定している。この特徴の一つはAIのライフサイクルを意識した設定をしていることであり、社会インフラのように長期的な運用が求められる事業に多く関わっている経験を踏まえ、日立では共通して求められる実践項目だけでなく、計画、社会実装、維持管理という各フェーズを意識して対応する行動基準を定めている。

次に、設定したゴールに向けてガバナンスが回るためのシステムデザインに関して、運用を回すための組織、人、仕組みの整備を行っている（図3参照）。

具体的には、Lumada Data Science Lab. (LDSL) というデータサイエンティストの組織の中に多様な専門人材で構成するAI倫理専門チームを設置するとともに、プライバシー保護諮問委員会など、日立内の他の関連組織との連携

体制を整備している。また、外部の有識者から構成されるAI倫理アドバイザリーボードを設置し、客観性を担保するとともに、ナレッジの蓄積を図っている。このAI倫理アドバイザリーボードとの連携も含めて専門チームは、いわゆるCoE（Center of Excellence）として機能しており、社内の情報共有や教育・人材育成という役割も担っている。

運用に関しては、AI倫理に関するリスクアセスメントを行っており、チェックリストや専門チームやアドバイザリーボードによる評価、それに基づく対策等を実施している。「AI倫理・バイ・デザイン」という考え方で、なるべく企画設計段階、あるいはPoC（Proof of Concept）の段階からゴールとなるAI倫理原則を評価し、対応するようにしている。

最後に、ループの中央の評価に関して、日々の運用だけでなく、半年に1度程度、このシステム自体を定期的に評価し、改善するということを行っている。そのためにも先ほど紹介したようにCoEである専門チームはAI倫理に関する政策や技術等、常に新しい情報を学習している。

講演（3）

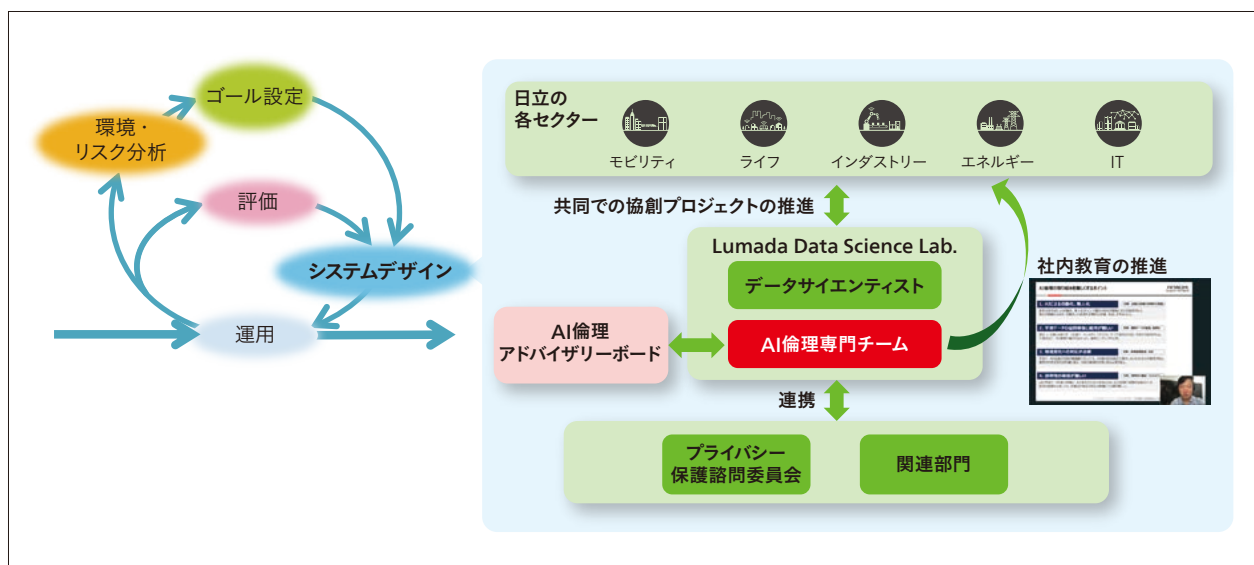
社会イノベーション事業のためのAI倫理

日立におけるAIの研究開発・社会実装（柳井 孝介）

日立の社会イノベーション事業は、健康・医療・介護から移動、金融まで、社会インフラや公共性の高いサービスに関わる。そのため、日立の事業にAIを活用していく際に

図3|システムデザインとしての組織、人、仕組み

専門組織の設置、人材育成、外部連携も含め、CoE（Center of Excellence）としてナレッジを集約している。



は、社会への影響が極めて大きいと自覚し、倫理観をもってAIを開発・活用していくことが必須である。これまで日立で実践してきたことを、もう少し具体的に述べたい。

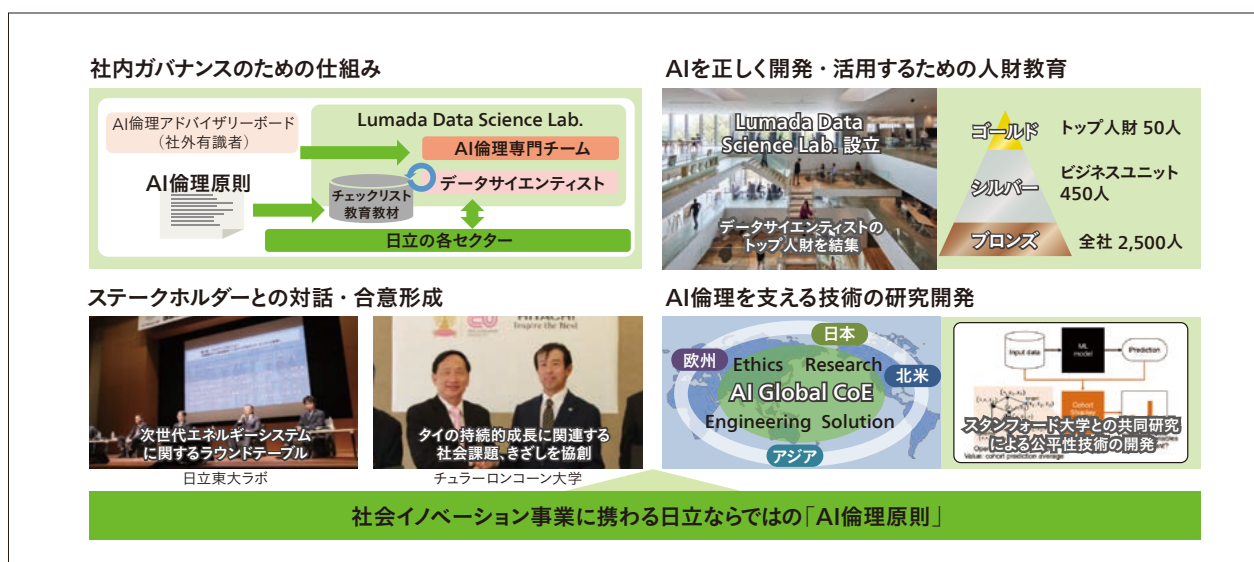
図4は、日立のテクノロジーガバナンスに関する取り組みの具体例を示したものである。詳細は別稿「社会イノベーション事業のための日立のAI倫理原則とその実践」（本号15ページ）を参照いただきたい。日立では、(1) 社内ガバナンスのための仕組みの構築、(2) AIを正しく開発・活用するための人材教育、(3) ステークホルダーとの対話・合

意形成、(4) AI倫理を支える技術の研究開発、などを重視して取り組んできている。またこれらの取り組みを推進するための共通の指針・言語として、社会イノベーション事業を意識した日立独自のAI倫理原則を定めている。

AIや機械学習などの新しい技術に関して、その技術的な特性や社会に与える変化などを深く理解しておかなければ、誤った利用をしてしまう恐れがある。AI倫理を実践するためには、まずAIそのものに関する深い理解が不可欠であり、長期的かつ組織的に取り組む必要がある。

図4|日立がめざすAI社会とそれに向けた取り組み

人がAIを信頼し、社会が合意の下、AIに多様なタスクを安心して任せられる世界をめざす。



日立はこれまで、東京大学に設置した日立東大ラボやタイのチューラーロンコーン大学とも連携しながら、ステークホルダーとの対話・合意形成に取り組んだ経験がある。AIの研究および社会実装を進めるにあたっては、マルチステークホルダー間での対話・協調が重要だと考えており、取り組みを強化していく。

AIの研究開発においては、日本、アジア、北米、欧州の日立のR&D (Research and Development)拠点をまたがって、「AI Global CoE」という枠組みを構築している。多様な視点を取り込み、包摂的に研究開発を進めている。AI倫理には、安全性、公平性、透明性、セキュリティ、プライバシーなど多様な観点があり、日立の研究の取り組みもさまざまである。例えば、スタンフォード大学とは、XAI (Explainable AI: 説明可能なAI) を用いて公平性を高める技術に関して、世界トップレベルの研究を共同で推進している。またAIそのものの研究以外にも、データのガバナンスの研究や、トラスト・ガバナンスの研究も推進している。具体事例の一つとして、XAIを活用して、救急体制の配備最適化を行うシステムを開発した例がある。公共性の高いサービスにおいては、高いレベルの公平性・説明性が求められる。そこで、公平性・説明性に配慮した技術を活用することで、社会実装を完遂することができた。

[セッション2]

ディスカッション

AIガバナンスとマルチステークホルダー間での協調

——AIガバナンス・ガイドラインの意義と企業での活用への期待に関して教えていただけますか。

泉 AIガバナンス・ガイドラインには、二つの意義があると考えています。一つは、各企業に対する支援です。AIの社会実装の促進に必要なAI原則の実践を支援すべく、事業者が実施すべき14の行動目標を提示し、それぞれの行動目標に対応する仮想的な実践例とAI原則からの乖離評価例を参考情報として提示しています。仮想的な実践例は具体的であり、ここで例示している乖離評価例も具体的であり、各企業がAIガバナンスを実践するうえで参考になると思います。もう一つは、AIシステムの開発・運用に関わる事業者等の取り引きなどで広く参照されることや、AI原則の実践に関するステークホルダーの共通認識の形成を通じて、各社の自主的な取り組みを後押しすることです。AIシステム利用者、つまり、エンドユーザーが実際にAIシステムを使うようになるまでに、多くのプレイヤーがAIシステ

ムの開発や運用に関与します。このバリューチェーンに関わるすべての企業が同じ目線でAIガバナンスを実践しなければ、負のインパクトが顕在化してしまうかもしれません。われわれはこのAIガバナンス・ガイドラインが多くのプレイヤーの目線を合わせるための共通言語になることを期待しています。AIガバナンスの議論と実践をリードする日立には、複数事業者間のAIガバナンスの実践のリーダーないしハブとなり、負のインパクトを軽減しながら、社会全体がAIから恩恵を受けられるように取り組んでもらえればと思います。

美馬 泉さんのおっしゃる通りAIガバナンスは一企業だけで達成できるものではありません。日立であればクライアントである企業がAIの運用者であり、AIを活用したビジネス全体としてAI倫理原則が遵守できるようにクライアントとのコミュニケーションを図ることも重要と考えます。

また、AIガバナンスに関する考え方については経済産業省で示していただいたように汎用性が高い部分も多く、企業をまたがって共用化していくことも不可欠です。日立は、AI倫理に取り組む他社とも個別に、あるいは業界団体を介して対話を行っており、連携を模索しています。

——AIの技術領域においては、アジャイル・ガバナンスが適しているとのことでした。このフレームワークが必要となる理由をもう少し詳しくお聞かせください。

泉 AIシステムに代表される、サイバー空間とフィジカル空間を高度に融合させるCPS (Cyber Physical System) を基盤とする社会は、複雑で変化が速く、予見可能性に欠け、リスクの統制が困難である場合が多いと思います。また、そうした社会の変化に応じて、ガバナンスがめざすゴールも常に変化していきます。したがって、常に変化する環境とゴールを踏まえ、最適な解決策を見直し続けるものです。そのためには、ゴールや手段が予め設定されている固定的なガバナンスモデルを適用することは妥当ではないと考えられます。最適な解決策を見直し続けるガバナンスの枠組みがアジャイル・ガバナンスなのです。特に、ガイドラインでは、ゴールからの乖離の評価と乖離への対応を必須プロセスとすべきとしています。この乖離評価にはAIシステムの開発や運用に直接関わっていない者が加わるようにすべきです。乖離があることだけを理由としてAIの開発・提供を不可とする対応は適当ではないと思います。乖離評価は改善のためのきっかけにすぎず、ループを回していくことが大切です。また、利用者に対して乖離の可能性や対応策に関する十分な情報を提供することも大切です。



柳井 実際に、AIの研究・社会実装を進めている立場からみても、そう思います。社会の状況の変化が速い、技術の進化スピードが速い、というのが根本なのですが、結局は社会から真に必要とされるもの、社会から信頼されるものをつくらなければ意味がありません。社会イノベーション事業を推進している日立がテクノロジーガバナンスに関して責任をもって進める必要があるのですが、政府機関や学術機関、市民・コミュニティなど、複数のステークホルダーと対話し、ステークホルダー間で信頼を構築しながら進める必要があります。そのような観点からもアジャイル・ガバナンスのループを回すことは重要だと思っています。

——AIのガバナンスを実現・実践するためには、さまざまなステークホルダー間で、具体的にどのような形で協調していくべきなのでしょう。

泉 ステークホルダー間で協調するためには、共通言語が必要です。その共通言語として、AIガバナンス・ガイドラインを作成しました。AIガバナンス・ガイドラインを活用して、社会的な共通認識を醸成していくことが重要です。ステークホルダー間の協調はさまざまであり、「one-size-fits-all」といった魔法の解決策はありません。それぞれが工夫していくことが不可欠だと考えています。

柳井 この点は手探りだと思います。ガバナンスも柔軟に継続的に迅速に、見直しを続けていくべきものです。ここでステークホルダーの多様な見識を集めることが必要です。日立はSociety 5.0の実現に向けて、これまでもステークホルダーとの対話・合意形成を意識して進めてきています。これらの経験を生かして、AIに関しても、より一層マルチステークホルダー間での協調という観点を意識して進めたいと思っています。まず企業側の第一歩としては、ス

テークホルダーに対するアカウンタビリティを今より一歩踏み込んで果たすことが重要なのではないかと考えています。

泉 多くの日本企業に共通して言えるのは、みずから動いてルールをつくるという動きが少ないということです。批判の矢面に立つことになることもありますが、めげずに活動することが大事です。日本企業にもぜひそういった動きを活発にやっていただければと思っています。みずから動く方が、早めにリスクに気づくこともあります。何らかの指摘を受けた場合でも、前向きにとらえて前向きに対処する姿勢を見せてもらえるとありがたいです。

——最後に改めて日立に対する期待をお聞かせください。

泉 日立には、AIガバナンスの実践のリーダーないしハブとして、ビジネスパートナーやエンドユーザーを含めたステークホルダー全体の協調・共通認識を先導してほしいと思っています。自社だけでは解決できない課題もあります。パートナー企業や顧客企業、他のセクターの人たちにも、みずからルールづくりに関与するつもりで意見交換などを積極的に行ってもらいたいですね。

柳井 おっしゃられる通りです。日立のAI倫理アドバイザリーボードの社外有識者の方々にも指導いただきながら、AIガバナンスの実践をリードできるよう尽力してまいります。またルールづくりに関しては、各コミュニティ・関係機関で意見を出すことに加えて、ステークホルダーが関与する実証実験に参加するなどして、社会が合意するデファクトスタンダードの策定にも貢献していきたいと思っています。

——本日はありがとうございました。