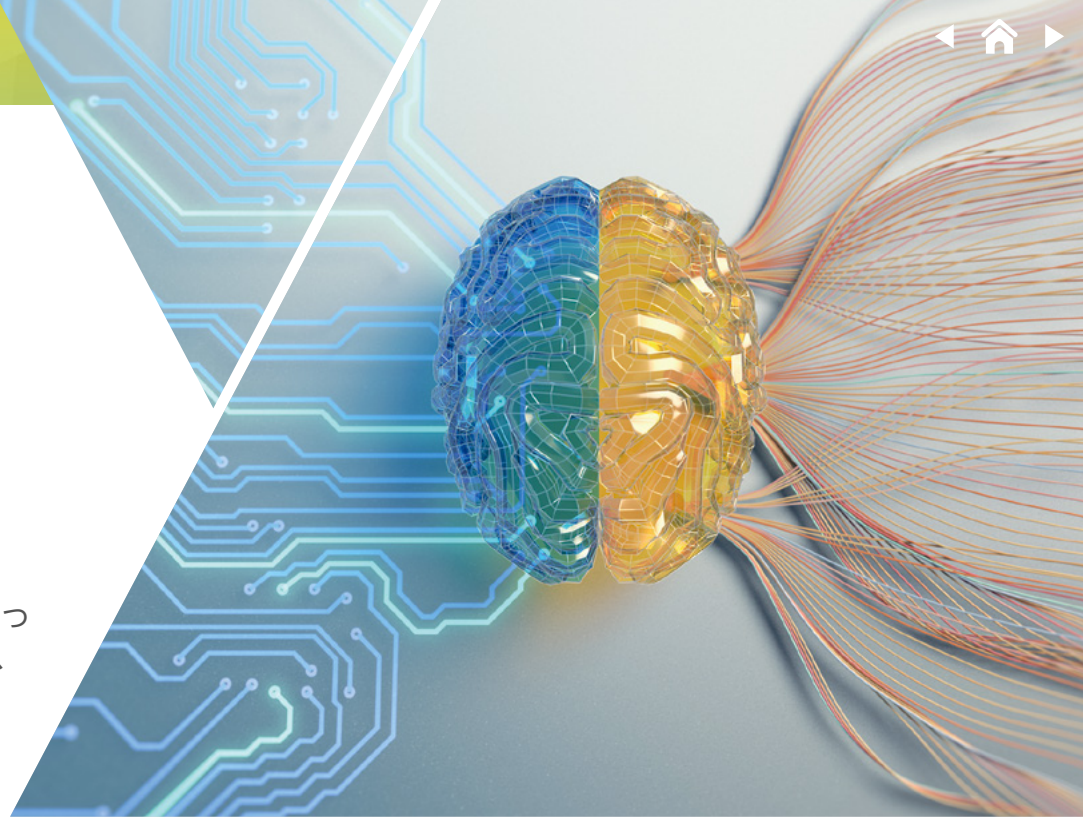
Part1

ICT 2023 トレンド

安心安全なAI社会に向けて、今考えるべき「AI TRiSM」とは？

AI技術の急速な進化・発展に伴い、いつの間にか私たちの社会にAIが浸透しつつある。より便利で豊かなAI社会への期待が高まる一方で、倫理や信頼性、セキュリティなどの観点から、AI活用に対する懸念も広がっている。今注目される「AI TRiSM」をキーワードに、AI活用に潜むリスクや、それらのリスクにどう向き合うべきか、富士通の取り組みとともに解説していこう。



取材協力・監修

富士通株式会社
AI倫理ガバナンス室

室長
荒堀 淳一

富士通株式会社
研究本部
AI倫理研究センター
プロジェクトディレクター
稲越 宏弥

※取材時の役職を記載しています。

暮らしに、ビジネスに広がり続けるAI活用

かつては「遠い未来の話」「SF小説の中の出来事」と思われていたAIが、近年、いつの間にか私たちの社会に広がりつつある。

例えば、日常生活に欠かせないスマートフォンにも、学習用からビジネス用、娯楽用まで、多くのAIアプリが搭載されている。また、ビジネスシーンにおいても、自動問い合わせ対応（チャットボット）や広告配信、顧客分析、採用マッチングなど、AI活用のフィールドが拡大し続けている。それと気付かぬままにAIを活用していることもあり、その意味ではすでに「AI社会」が到来しているとも言える（図1参照）。

では、そもそもAIとはどんな仕組みで動いているのだろうか。AI(Artificial Intelligence：人工知能)とは、その名のとおり、人の知能を人工的なコンピューターによって模倣するもの。大量のデータから判断基準やより一般的な概念を学習し、人間を上回るスピード（計算速度）で結果を導くことが可能だ。少子高齢化が進む先進国では減少する労働人口を補うために、また、技術者が不足する途上国では社会インフラや医療などを支えるために、今後もAI活用への期待は高まり続けることが予想される。

AI倫理ガバナンス室の荒堀氏が「AIは健康や福祉、気候変動の抑止、平等社会の実現などSDGsの達成へ

の貢献も期待され、ひいては人間社会の発展と進化にも貢献できる素晴らしい道具」と語るように、富士通はAIを社会課題の解決に向けたキーテクノロジーの1つに掲げ、開発・適用を強化。2019年度以降で実に5,000件以上のAIサービスを国内外に提供している。

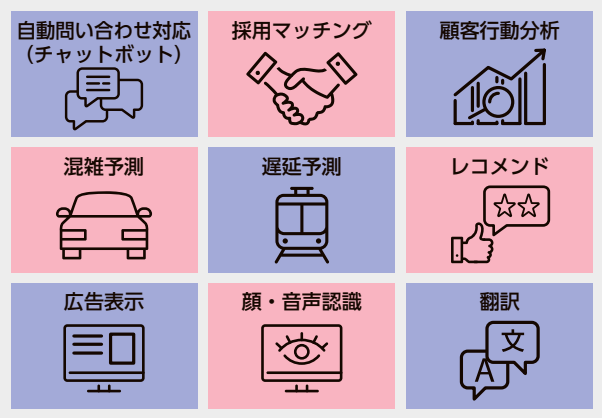


図1）広がり続けるAI活用のフィールド

世界的に注目が高まるキーワード「AI TRiSM」とは

AI活用が社会に広がる一方で、「AIの判断を無制限に信用しても良いのだろうか」といった懸念の声も高まっている。

AI倫理研究センターでプロジェクトディレクターを務める稲越氏は、AIを巡る世界的な潮流について、次のように語る。「AI活用によるイノベーションが世界規模で加速する一方、顔認識AIが人種差別的な結果を出したり、採用マッチングAIが性差別的な結果を出したり、あるいはチャットボットが差別用語を連発したりと、様々な問題が報道されています。加えて、AI活用による公共監視やプライバシー侵害、世論誘導、さらには一部の巨大IT企業による寡占状態への懸念もあり、AI活用に対する規制強化への議論が活発化しています」。

こうした背景のもと、近年、注目を集めているキーワードが「AI TRiSM」だ。米国の調査会社ガートナーが発表した造語で、「2023年の戦略テクノロジーのトップトレンド」の1つにも挙がっている。

AI TRiSMとは「AI Trust, Risk and Security Management」の略で、「AI活用に伴うリスクに対応し、AIの信頼性を高めるための取り組み」を意味する。具体的には、AI活用における倫理性や透明性、説明可能性、外部からの攻撃に対するセキュリティ、プライバシー保護、適切なマネジメントやガバナンス、それらをサポートする手法やツール、プロセスなど、広範なテーマが含まれている。

これからAIを開発・利用しようとする企業にとっては、理解しておくべきキーワードと言えるだろう。

なぜAI活用にリスクが伴うのか

AI活用に伴うリスクは、どのような原因で生じるのだろう。荒堀氏は、次のように説明する。「AIは、大量のデータを学習することにより、共通セグメントの一般的な傾向は算出できるものの、対象となった一人ひとりの個性は切り捨てられる場合もあります。また、AIに学習させるデータにバイアス（偏り）があると、AIが導き出す結果にも偏見や差別的バイアスが生じてしまうリスクがあるのです。こうした特性を理解せず、AIの判断を鵜呑みにしてしまうと、バイアスを是正できないまま、つまり不公平や不平等を放置したまま意思決定をすることになりかねません」。

こうした技術・学習データ面の問題に加え、AIの「ブラックボックス化」も大きな課題となっている。理由や根拠を客観的に示すことができないAIに、人の生

命や財産に関わる重要な判断を任せられないと感じるのは、ある意味で当然の心理だろう。例えば、最終的な判断は必ず人間が下すなど、リスクを排除する仕組みを構築することが、今後のAIの信頼性を大きく左右するだろう。

では、これからAIを開発・活用しようとする企業には、どのような姿勢が求められるのだろうか。ここで強調しておきたいことは2つある。1つは、荒堀氏が言う「AIの特性やリスク」を理解すること。もう1つが「AI倫理」に取り組むこと。「AIを開発・提供する企業や、利用する企業の経営者は、AI倫理ガバナンスを推進して、倫理的な観点の必要性を従業員にしっかりと浸透させていく必要があります。従業員全員がAIのメリットとデメリットを理解したうえで、ユーザー企業様もAIの本質を理解し、AIが不都合をもたらす可能性を認識して利活用しなければならないと思います」（荒堀氏）。

富士通が取り組むAI倫理

富士通は、1985年にLisp言語高速処理用専用コンピューターを商品化するなど、30年以上にわたり日本のAI研究をリードし続けてきた。こうしたAI分野におけるリーディング企業としての責任から、「AI倫理」にもいち早く取り組んでいる。その牽引役として、2022年2月に設立されたのが、荒堀氏が室長を務める「AI倫理ガバナンス室」だ。

「私たちは、『イノベーションによって社会に信頼をもたらし、世界をより持続可能にしていく』というパーパスを掲げています。AIをはじめとする最新のキーテクノロジーによって人を幸せにするために、AI倫理がきちんと担保された、『信頼できるAI』を提供する責務があります」と荒堀氏が語るように、AI倫理は富士通のパーパスと密接に結びついており、2019年3月に策定・公表されたAI倫理指針「富士通グループAIコミットメント」も「私たちのパーパスをAIの観点から具体化したもの」だという（図2参照）。

このコミットメントのもと、全社的なガバナンスや社外も含めた意識浸透を担うAI倫理ガバナンス室と、国内外の研究機関と連携しながらAI倫理に関する先進技術を研究する「AI倫理研究センター」および「人工知能研究所」が、いわば社会的アプローチと研究アプローチの両軸から取り組んでいる。

また、これらの取り組みを支えるべく、2019年9月から「富士通グループAI倫理外部委員会」を開催している。この委員会は、国内外の社会情勢を踏まえ、ステークホルダーから客観的な評価を受けることを目的に、約半年に1回、多様な分野から専門家を委員として招聘し、AI倫理について多様なテーマで議論している。「AI活用により影響を受ける人々の人権をどう守るか」「AIリスクの判断を現場任せにしない仕組みづくり」「社外連携の重要性」などが議論されており、富士通の体制強化や施策に反映されている。

トラステッドなAI社会の実現へ。
倫理的課題にどう向き合うか

富士通グループ AIコミットメント5項目

1. AIによってお客様と社会に価値を提供します
2. 人を中心に考えたAIを目指します
3. AIで持続可能な社会を目指します
4. 人の意思決定を尊重し支援するAIを目指します
5. 企業の社会的責任としてのAIの透明性と説明責任を重視します

図2）富士通グループAIコミットメント5項目

「AI倫理影響評価技術」を開発し、無償で公開

AI倫理を原則から実践へと推し進めていくため、富士通ではAI倫理研究センターを中心に、技術的観点からも様々な取り組みを行っている。

「富士通では、AIの信頼性を高めるためには、AIを法や規範に準拠させるよう、その手順を定めるだけでは不十分だと考えています。より重要なのは、法や規範を逸脱するリスクを査定すること。つまり、AIを悪用しようとする人が出るかもしれない、意図せずに社会に悪影響を与えてしまうかもしれない、といった潜在的なリスクを評価することです」（稲越氏）。こうした考えのもとに開発されたのが、AI活用に伴う倫理的な影響を評価し、どのようなシーンで問題が生じるのかを洗い出す「AI倫理影響評価方式」だ。富士通は「信頼できるAI」を社会に普及させるため、この方式に基づく手順書や適用例を、2022年2月からAIシステムの開発者や運用者向けに無償公開している（コラム参照）。

安心安全で豊かなAI社会の実現に向けて

AI倫理やセキュリティリスク、プライバシー侵害などAI TRiSMへの要求は今後も高まり続けるだろう。無条件・無制限な利用に対する規制がブレーキになるかもしれないが、だからと言ってAI活用の潮流がとどまるとは考えづらい。

「近年、AI活用のハードルは大きく下がってきています。かつては解きたい問題をAIに当てはめるために高度な専門知識・技術を必要としましたが、そうした知識・技術を備えた研究陣の努力の甲斐あって、今や専門知識がなくともAIを活用できる時代が到来しつつあるのです。実際、ここ数年で汎用性のある「基盤モデル（foundation model）」が登場し、大量のデータでゼロから学習する従来の手法とは対照的に、少量のデータで個々の問題に対応できるようカスタマイズする手法が主流となっています」と稲越氏は語る。これに伴い、AI技術者に求められるスキルも変化しているという。「今後も専門知識・技術が必要であることに変わりはありませんが、加えて、AIをいかに使いこなすか、いかに社会に受け入れられるようにするかも問われていくでしょう」（稲越氏）。

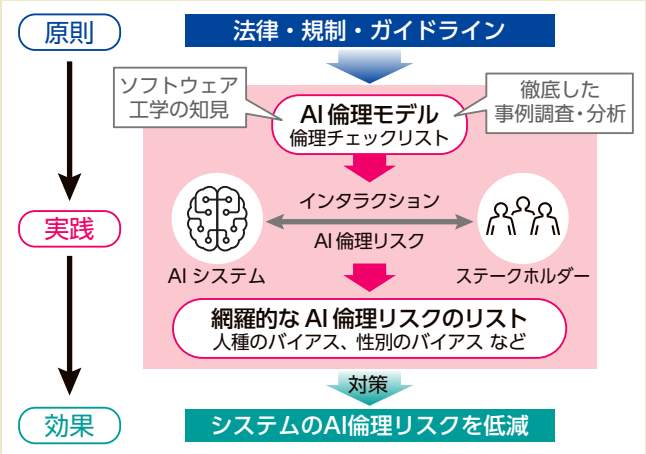
一方で、荒堀氏は次のように語る。「技術革新によってAI活用の可能性は広がり続けています。センサー技術の発展で今まで測れなかったデータを測定できるようになり、コンピューター技術の進化で膨大な計算が瞬時に行えるようになりました。これからのビジネスパーソンに求められるのは、今まで見えなかったものが瞬時に見渡せるようになった世界を想像できるかどうか。その力を持った人材の有無が企業の競争力に直結します」。

AIをビジネスに活かすためには、「AIは善か悪か／必要か不要か」といった二元論でなく、個々のユースケースをしっかりと具体的にイメージし、そこで生じうるリスクや、回避するためのルールを考える姿勢・思考が求められる。そうした積み重ねの先に、人とAIが共生する、安心安全で豊かなAI社会が実現するのではないだろうか。

column

AIによる倫理影響をいかに評価するのか

富士通では、ソフトウェア工学の知見と徹底した事例調査から、AI倫理に関するリスクが、AIとステークホルダー（開発者や運用者、ユーザーなどの関係者）とのインタラクション（相互作用：情報のやり取りなど）に由来することを着想。2019年前後に編纂された倫理原則やガイドラインと、過去にAIが引き起こしたインシデントを照らし合わせて、AIシステムにおいて考慮すべき項目を体系化した「AI倫理モデル」を開発し、AIシステムにおける倫理リスクの系統的かつ網羅的な評価を可能にしている。



富士通が開発したAI倫理影響評価技術

Part2

ICTトレンド2023連動企画

特別座談会 | 会員企業若手社員×富士通有識者

ビジネスパーソン一人ひとりが考えるべき AI倫理とAI社会の未来

座談会参加者　荒堀 淳一／稲越 宏弥

株式会社マルハン
グループユニット
経営企画部情報システム課
梁 麗麗 氏

FITEC株式会社
ビジネスソリューション事業部
事業IT・デジタル戦略室
イノベーション推進課
野村 遼平 氏

日本通運株式会社
IT推進部
大川 拓人 氏

AIが社会にもたらす価値やリスクは、AIを開発する側、利用する側を問わず、これからのAI社会を生きる一人ひとりが考えるべき問題です。そこで本記事では、FUJITSUファミリ会会員企業の若手社員3名にお集まりいただき、ICTトレンド記事の監修を務めた富士通有識者に対して、それぞれが所属する企業や担当業務などを踏まえた、AI活用やAI倫理に対する疑問や考えをぶつけてもらいました。様々な視点からの意見が交わされた議論から、AI社会の未来を考えるうえでのヒントを見つけてください。

そもそも、AI倫理とは何を対象にしているのか

梁　AIを活用するうえで「AI倫理」について考える必要があると言われていますが、正直に言うとイメージしづらい面もあります。そもそもAI倫理とは、AIそのものを対象としたものでしょうか。それとも開発者やユーザーを対象としたものでしょうか。

稲越　おっしゃるように、一口に「AI倫理」と言っても いろいろな意味があります。大別すると2つの視点があり、1つはAIそのものの人権を考える視点。もう1つがAIを扱ううえで倫理に反するような使い方になっていないかを考える視点。現在、世界的に問われているのは後者であって、富士通もそこをテーマにしています。

野村　私もAI倫理が誰を対象にしているのか疑問でしたので、はじめに概念が整理されたのはありがたいですね。

梁　素朴な疑問ですが、なぜAI倫理を考える必要があるのでしょうか。

稲越　AIは人間の脳の働きをモデル化したものですが、実際の原理は異なっています。ですから、人間の作業をAIに代替させると、人が当たり前に持っている倫理観に反した判断を下しかねません。例えば「この人にはお金を貸さない」「この人は採用しない」といった、いわば“血も涙もない”対応をしかねないリスクがあります。このため、AIを人間社会で活用する際には、人が持っている倫理観に反する行動をしないかどうかを考えていく必要があります。

大川　人間とAIとで原理が異なるというのは、どういうことでしょう。

荒堀　ビジネスの世界で例えれば、企業の売上が下

がった際、私たち人間はすぐに値下げするのでなく、なぜそうなったのか理由を分析し、原因を究明したうえで対策を練りますよね。ところが、AIは因果関係よりも相関関係で考えがちなので、実際には売上低下に関係ない部分を改善するなど、間違った結果を導きかねないリスクがあります。

大川　因果関係と相関関係を混同するのは、AIだけでなく私たちもやってしまいがちです。AIを利用する人間側も、両者の違いをしっかり認識する必要がありますね。

荒堀　AIは大量のデータを学習して、そこから人間の判断基準を学習しますが、学習させるために収集したデータの中に誤ったデータが含まれていたり、データ自体に偏りがあったりすると、それに引きずられてAIが間違った判断を下しかねません。AIの判断を鵜呑みにするのではなく、こうしたAIの特性をしっかり把握したうえで利用する必要があります。

梁　AIが誤った判断を下すのは、具体的にはどんな場合が考えられますか。

稲越　例えば、収集したデータがブログやツイートなどSNSからの情報に偏った場合、過激な意見や誇張した意見など、個人の偏見やバイアスがかかったデータばかりを集めてしまい、AIの判断も引きずられてしまう恐れがあります。

荒堀　ひとつ具体例を挙げると、ある企業が採用マッチングにAIを活用した際、過去の採用データをもとにしたところ、それまで男性の採用が中心だったために「この会社は男性を採用すべき」と誤った判断を下し、女性が採用されないといった事態を招きかねません。

梁　なるほど、偏りなくデータを集めたり、AIが下した判断を検証したり、AIを活用する側の配慮が大切だということが理解できました。

AI倫理にまつわる法整備はAI活用にブレーキをかけるのか

野村　最近ではイラストAIと著作権の問題が取り沙汰されるなど、新しい利用法が生まれると、あとから法規制の問題が浮上して、活用する側にとってはリスクになります。AIによるイノベーションと法規制リスクとのバランスをどう考えるべきでしょうか。

荒堀　もともと法律というものは、倫理や道徳だけでは対応できない問題が生じた際に、必要に応じて作られるものなので、常に「後追い」になります。例えば、今、例に挙がった著作権法はクリエイターの権利を守るために作られたものですが、一方で、自由競争を守るために独占禁止法という別の法律があり、全体でバランスを取っています。法規制リスクについても同様に、これから様々な状況に応じて法整備が進んでいくと思われます。

野村　法規制に至る社会の反応をしっかりと見据えながら、その時々でバランスを取っていくことが大切ですね。

荒堀　AIに限らず、新しい技術が登場すると、必ずそれらを悪用しようとする人が出てくるのは避けられません。一方で、正しく使おうとしても、うっかり倫理に反する事態を招く恐れもあります。これらは別々に考える必要がありますが、いずれにしても新しい技術に対応したルールが必要になることは間違いありません。とはいえ、現時点での課題の多くは既存の法律で解決できるので、1つひとつ調べて判断していけば、リスクを排除できるはずです。

稲越　先ほどのイラストAIで言えば、問題ごとにそれぞれ考える必要があります。例えば著作権はAIにあるのか、それともAIを開発した人にあるのか。また画像を収集するのは認められるのか。商用利用は不可で、実験・研究用なら可など、同じ行為でも目的によって判断が変わるなど、解釈が難しい面がありますが、ユースケースごとに問題点を明確にして、丁寧に洗い出していくことをおすすめします。

大川　私は物流会社で海運システムを担当していますが、システム開発中に新しい法律が整備され、システム改修が必要になったことがあります。新しい分野で、いつ新しいルールができるかわからないというのは開

発上のリスクになります。

野村　同感です。「どんなリスクがあるかわからない」という不安が先立ち、どうしてもAI活用にブレーキがかかってしまうのではないでしょうか。

稲越　AIによる利便性とリスク、どちらを重視するかは難しい問題で、世界的に見ても意見が分かれています。EUなどは「ルールを定めたうえで慎重に活用すべき」という態度ですが、米国や中国では「イノベーションを阻害しないよう、あらかじめ規制するよりも、まずは積極的に活用して、問題が出ればあとからルールを決めよう」というスタンスの違いがあります。日本はどちらかと言えば後者で、ガイドラインなどはあっても、明文化されたルールは定められていません。

梁　そうしたスタンスの違いは、どちらが良い悪いという問題ではなく、国民性や政府の方針による違いと考えて良いものでしょうか。

荒堀　おっしゃるように、イノベーションを優先する米中も、慎重派の欧州も、根っこにあるのはAIで社会や生活を豊かにしたいという思い。それはAIを開発する私たちにも共通するものです。

稲越　どちらが正しいと断言できるものではないので、どうバランスを取るかという質問に対しては、「これが正解」という答えはありません。それでも常に考え続けなければならない重要な問題だと思っています。

グローバル社会で生じる国や地域によるAIへの温度差

大川　例えば自動車メーカーでは、ガソリン自動車が求められている市場がある一方で、EV以外は禁止という市場もあるなど、地域ごとのギャップをどう埋めるかが課題になっています。AIについても、先ほど言われたように欧州と米中での温度差がある中、どう対応すべきでしょうか。

荒堀　厄介なことに、法律などのルールは、文化や商習慣、政治体制などを背景に定められるので、国や地域によって差が出る場合があります。目的は同じでも、方針や解釈が異なるところに難しさがあるので、グローバルとローカルの両面で考える必要があります。

野村　私は社内でITツールを推進する役割を担っていますが、同じ社内でも部署によって反応が異なることがあります。国や文化が異なると、違いはより大きくなるでしょうから、リスクと利便性のバランスを取るのは非常に難しくなりますね。

稲越　エネルギーの話題で言うと、ある地域で代替エネルギー活用が進んで炭酸ガスの排出を抑制できたものの、その影響で他の地域では大幅に排出量が増えたというデータがあります。AIについても、その活用による影響を、自国だけでなくグローバルで考える必要が出てくるでしょう。

梁　最近注目を集めている「ChatGPT」のように、1つの企業、1つの国・地域での利用にとどまらない活用事例も増えています。それらを使った際の責任について、世界規模で考える必要が出てきますね。

荒堀　グローバル社会では、ある地域で生まれた新技術が、瞬く間に世界に広がっていきます。バックグラウンドの異なる多くの人が一緒に考えていくべき問題だと思います。

大川　AIの活用法も多種多様で、社会課題の解決に役立てようとする方向もあれば、日常生活を豊かにしようという方向もあります。目の前の利益に目を向けがちですが、それでいいのかという意識もあります。そうした意識を持つ人が増えていくことが大切ではないでしょうか。

稲越　AIが技術面だけでなく産業として成熟しつつある中、社会に対する影響力も強まっています。AIを社会に調和させていくために、事前にAI活用のリスクを考えるとともに、AI活用のあり方について広く社会全体で考えていけるよう、積極的に情報発信していきたいと思っています。

求められるのは対話を通じて着地点を導く姿勢

野村　地域や時代背景に加えて、業種業態によっても、AI活用のリスクの捉え方は変わってくると思います。当社グループのような製造業、マニュファクチャリングにおけるAI活用のリスクはどう捉えれば良いでしょうか。

荒堀　医療／ヘルスケアや自動車の自動運転のように、人の生命・安全と直結する分野に比べればリスクは少ないと言えますが、だからと言ってリスクがないとは言えません。やはり様々なケースを想定して、個別に考える必要があります。

梁　私たち娯楽サービス業でも、サービスの質や、店内での導線、顧客分析など、AIを様々に利用できます。そうした用途ごとに、リスクや対策を考えるべきということですね。

稲越　AI活用という大きな括りで考えるのではなく、ユースケースごとに論点を明確にして考えることが大切です。ビジネス現場の方々と対話していても、AI倫理にピンとこない方でも、セキュリティやコンプライアンスなど具体的なテーマに落とし込めば、議論が具体的になるものです。

大川　難しいことですが、そうやって論点を明確にしながら、対話の中で着地点を見つけていくという姿勢が大切ですね。

荒堀　そのとおりです。AIは人間よりも速く大量のデータを処理して私たちをサポートしてくれます。間違いなく社会を豊かにしてくれるものですから、世界で話し合いの場を設けて、価値観を共有することが重要です。5月に開催されるG7広島サミットでも、AI倫理についての議論が期待されています。

座談会を振り返って

梁　AIは近いようで遠い存在と思っていましたが、今日の対談を通じて、より身近に感じるとともに、興味のあるところから自分で調べて、そのメリットやデメリットを体感するのが大切だと感じました。

大川　リスクに配慮して何もしないとイノベーションは起こせないので、チャレンジする重要性を感じ、後押ししてもらえた気持ちです。今日の経験をこれからの仕事に活かしていきたいと思います。

野村　AIに対していろんな考えがあっても、社会を豊かにしようとする思いは共通しているというご意見が印象に残っています。AIに対する意識や意見の異なる相手とも、しっかり話し合うことが大切だと感じました。

稲越　多様な視点からご質問やご意見をいただけたおかげで、こちらも考えを整理でき、対話する意義や大切さを改めて感じました。今日の学びを今後の研究に活かしていきたいと思います。

荒堀　AIは歴史が浅い分野だけに、皆様のような若い人たちが最先端の動向や背景を丹念に追っていけば、よりAIを使いこなせるようになるはずです。AIを世の中に役立てていくためには、多様な視点、多様な考え方が必要ですが、それらは多くの人の話を聞くことで培えるもの。今後もいろいろな人と話をして、いろいろな角度から考えてもらいたいと思います。そのためにも人脈を広げて、多様な背景や価値観を持った人と対話できるといいですね、今日の場がそのひとつになれば嬉しく思います。