消費者委員会本会議（第475回）
議事録

内閣府消費者委員会事務局

消費者委員会本会議（第475回）
議事次第

１．日時　令和7年11月19日（水）　10時00分～11時22分

２．場所　消費者委員会会議室及びテレビ会議

３．出席者

（委員）

【会議室】　　鹿野委員長、黒木委員長代理、中田委員

【テレビ会議】今村委員、大澤委員、小野委員、柿沼委員、善如委員
原田委員、山本委員

（説明者）

国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー
妹尾様

（事務局）

小林事務局長、吉田審議官、友行参事官

４．議事

(1)開　会

(2)AI・デジタル技術の利用により生じうる消費者問題について（有識者ヒアリング）

(3)閉　会

《１．　開会》

○鹿野委員長　本日はお忙しいところ、お集まりいただきありがとうございます。

ただいまから、第475回「消費者委員会本会議」を開催いたします。

本日は、黒木委員長代理、中田委員、そして私、鹿野が会議室にて出席しており、今村委員、大澤委員、小野委員、柿沼委員、善如委員、原田委員、山本委員がテレビ会議システムにて御出席です。なお、一部の委員は遅れての御参加と伺っております。

それでは、本日の会議の進め方等について、事務局より御説明をお願いします。

○友行参事官　本日もテレビ会議システムを活用して進行いたします。

配付資料は、議事次第に記載のとおりでございます。もしお手元の資料に不足などがございましたら、事務局までお申し出くださいますようお願いいたします。

以上です。

○鹿野委員長　ありがとうございました。

《２．　AI･デジタル技術の利用により生じうる消費者問題について(有識者ヒアリング)》

○鹿野委員長　本日の議題は、「AI・デジタル技術の利用により生じうる消費者問題について」です。

AI・デジタル技術の進展は、消費者の利益の増進に資する側面が大きい一方で、新たなリスクや課題を生じさせるおそれもあります。第8次消費者委員会で取りまとめられた「次期消費者委員会への移行に当たっての留意事項」においても、「デジタル化・AIへの対応については生じうる消費者問題を想定し、あらかじめ対応を講ずることが重要」であるとされているところでございます。

そこで、本日は有識者として国立研究開発法人産業技術総合研究所の妹尾様より、特に生成AIの発展という観点から消費者に関わる課題について御発表をいただき、意見交換を行いたいと思います。

改めて御紹介させていただきます。本日は、先ほども申しましたが、国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサーでいらっしゃいます妹尾様にオンラインにて御出席をいただいております。本日は大変お忙しいところ、どうもありがとうございます。

それでは、20分程度で御説明をお願いします。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　妹尾でございます。

それでは、御発表させていただきます。AIセーフティーとリスクマネジメントという形

についてというところで、消費者視点での課題と今後の問題について御説明させていただきたいと思います。

略歴は飛ばしまして、後で見ておいてください。

内容ですけれども、AIの制御というか、AIをどうコントロールしていくかというものに関する世界の動向、それから、課題の本質は何か、それから、生成AIについて今、世界としてどんな取組をされているのかの4つです。それから、生成AIの品質マネジメントとはどういうものになっているのかというのを御説明して、最後にというか、こちらが本題ですけれども、生成AIの発展と消費者に関わる課題ということで説明させていただきます。

基本的に、まずAIになぜ品質が必要かということなのですけれども、今はソフトウエアによる複雑な制御に人命を預ける時代ということで、今は世の中のほとんど全てのものにソフトウエアというものが使われて、それが人々の安全・安心に大きな影響を与えているというところで、そのソフトウエアにAIがかなり使われるようになってきているというところで、さらにソフトウエアに加えてAIのコントロールというのが重要になっているわけです。世界的に2019年頃からこういうAIに対する社会からの要求を明文化しようという動きがあって、OECDでAIのプリンシプルが2019年に策定されたり、日本でもそれに先駆けて人間中心のAI社会原則というところで、あくまでもAIを使うに当たって人間が中心なのだよとか、教育リテラシー、プライバシー、セキュリティー、公平性等々のポリシーが決められています。

社会から見たAIということで、更にAIがいろいろ発展していく中で、2020年代に入って法律やポリシー、それから種々のガイドラインというのがいろいろ策定されていて、米国はNISTが作っていますし、欧州が欧州委員会がAI法というのを策定中ですし、日本も経済産業省さん、総務省さん、内閣府のいろいろな活動があって、今はAI事業者ガイドラインというのが出ていて、AI基本法案も検討中という状況でございます。

欧州AI法ですけれども、AIに対して絶対こういう用途に使ってはいけないというリスクベースアプローチを取っているのですけれども、Unacceptable Riskということでサブリミナルやソーシャルスコアリングなど、人を何らかの格好で評定するという用途は絶対駄目よと。今、High Riskというところで医療応用などいろいろなところに対してAIがどうコントロールすべきかという規定が欧州整合標準という形でまとめられていて、来年8月に施行される予定ですけれども、ちょっと遅れ気味という話も聞こえてきています。

日本はAI事業者ガイドラインということで先ほどの人間中心のAI社会原則、それから経産省、総務省でいろいろなガイドラインがあったのですけれども、それらを一本化して2024年に事業者ガイドラインというのがまとめられていて、理念からポリシー、原則、それからどのようにコントロールすべきか、特にマネジメントレイヤーのガイドラインという形で策定しています。こういう安全・安心から始まって、公平性だとか、いろいろなものが記載されています。

AI規制についての大まかな動向なのですけれども、中国はかなり自由というか、国家が

それなりの統制はするのだけれども、基本的に自由にどんどんAIは使ってよい、それが今後の国家の競争力の源泉になるという考え方で運営されていて、EUはどちらかというとAIはかなり規制しないといけないと。GDPRからの流れもあって規制をかけるという路線で、米国、日本は割と中道で、必要最小限の制限はかけるけれどもできるだけ自由に使わせようということだったのですけれども、最近、トランプ政権に変わって米国がかなり自由主義のほうに行っている。それから、EUも中国、米国の状況を見ながら自由な方向に振っている。日本は閣議決定で人工知能関連技術の研究開発及び活用の推進に関する法律案というのでこういうものができて、必要最小限だけれども社会がおかしいことにならないようにというたがだけをはめて、あとはイノベーションと両立させるという方向ですけれども、どちらかというと日本も今後、更に自由な方向に行くのではないかと見られています。

機械学習AI、AIはなぜソフトウエアと違ってそういうものが難しいのという本質だけ御説明しますけれども、従来のソフトウエアというのは分析的アプローチで、要件定義して、設計して、実装して、それは人間がこのように動くというのを全部分析的につくり込むので、そのつくったものがちゃんと思ったとおり動いているかというのをV字型モデルといいますけれども、それらを検証していけばいいという話になっているのですけれども、機械学習の場合は要件定義があって、基本的には過去の経験というのをデータとして集めてきて、そこから過去の経験を数学的に表現するモデルをつくって、その過去の経験を使って将来を予測したり、判別したり、いろいろなことをするということで、要件定義、何かをやらなくてはいけないとなった後で、それに対して経験を集めるためにどういうデータを集めてどう勉強すればいいのかというのを決めて、データが機能をつくり込んでいくわけですね。なので、従来の方法論が通用しない。

我々は猫問題と呼んでいますけれども、例えば猫を判別するプログラムをつくれといったときに、適当に猫の写真やイラストなどのいろいろなものを集めてきて、それらを勉強させてモデルをつくる。そのモデルを使えば、割と簡単に猫の判別プログラムというのを、従来のソフトウエアは非常に難しかったのだけれども、そういうものが非常に簡単につくれますよということになったのですけれども、実際はその品質をきちんと考えようとすると、猫というのはどんな種類があるの、猫のどういうところが写っているの、尻尾だけとか、あるいはイラストの猫とか、いろいろな要件があって、これらのどういうイラストを、あるいはどういう猫を分類すべきかというところはビジネス応用で、本当に何に使うかにすごく依存するのですね。それらをきっちり定義して猫とはこういうものよというのをきっちり規定できればちゃんとした品質管理ができるのですけれども、それだとプログラムを書いてしまえばいい、AIの利点が消えてしまうのですね。そうすると、猫というのを適当に、少なくともこういう種類の猫とか、こういう部分とか、いろいろな限定を要件として決めて、その要件を満たす、それらをちゃんと分類できるためのデータを集めてきてモデルをつくってそれらを使うという風にしなければいけなくて、そういう意味では機能のつくり込みをデータがするというところで、しかもそれをとことんまで突き詰めて明確に

は記載できないというところで難しさがあるわけです。

それらをどのようにマネージするかということで、我々は2018年からずっと研究開発をやっていて、2020年に機械学習品質マネジメントガイドラインというのをまとめています。生成AIが登場して、更にAIに対してきちんと対応しないと非常に社会としては変なことが起こるということで、日本でも岸田元総理大臣が広島AIプロセスというのを広島のサミットで提唱されて、世界でこういうものを連携していこうという枠組みの中で日本は2024年にAIセーフティー・インスティテュートというのをつくって、世界も同等のいろいろな組織をつくってみんなで連携しようという話が動いています。

日本のAISIは、内閣府中心で12府省5関係機関ということで、省庁が主導して技術的なところを4つの国立研究機関がサポートするという体制になっています。我々は生成AIに対するマネジメントガイドラインというのも作っていて、生成AIというのは今は残念ながらほとんどアメリカのベンダーがつくったものを日本が使っているという状況になっていて、そういう意味では生成AIの大規模基盤モデル、LLMと言われる本丸のところになかなか日本のコントロールが効かないのですね。というところで、そういうものを使ってソリューションをつくり込むときにどういうところに気をつけなさいよというのを記載されたものになっています。

ここから、消費者に関わる課題について御説明いたします。特にサービスの開発者や提供者だけでは解決できずに消費者が一緒になって解かないといけない問題というのを挙げたつもりです。

1つ目が生成AIの生成物とフェイク、著作権の問題というところで、今は簡単に有名人のフェイク動画というのが作れる時代になっていますし、あるいは肖像権や著作権の保護という問題も非常に大きな問題になっています。今、こちらはいろいろなところで議論が進んでいますし、これは技術というよりは、技術で電子透かしを入れるとか、本物と偽物を区別するというところはありますけれども、最終的には社会のところでこれらに対してどう対応するのかという使い方という観点で規制をはめていくしかないような問題と認識しています。

2つ目が生成AIと教育界ということで、今、教育の現場もレポートの問題を出しても生成AIを使うと非常に簡単にレポートの構築というのができてしまうということで、あるいは入試も論文入試みたいなものですと、その論文の生成というのが生成AIを使うと非常に簡単にできてしまうというのがあって、論文入試は難しいとか、いろいろな先生方についてもどのようにレポート課題などの課題を生徒に出すか、実際にどう勉強していただくかというところはかなり頭を悩まされておられます。

それから3つ目、私はここが一番ポイントだと思っているのですけれども、先ほども言いましたとおり、AIというのは完全に品質を100％つくり込むというのが非常に難しい。AIの基盤となっているのが確率統計モデルというものをベースにしてつくられているというのが一つ。それから先ほど申しましたとおり、100％ちゃんと品質を管理しようとすると機能

を全て洗い出す必要があって、基本的にAIをつくって使うというところでそれは非常に現実的ではないという問題があって、そうすると、どうしてもあるところ以上の残存リスクというのが残ってしまう。その残存リスクに対して、当然提供者が取れればいいのですけれども、100%取るのが難しいとなったときに、それらを提供者と消費者とで分担する必要があって、そこをどうするかという課題があります。

もう一つは、生成AIの非常に高いいろいろな機能が実現してきていて、それが通常に使われてちゃんともともと設定したとおりに動いてはいるのだけれども、非常に強力な機能が社会との不整合を起こすという問題があります。

それからもう一つは、生成AIがどんどん発展していく中で、消費者、あるいは社会制度がそれに追随するのだけれども、この発展のスピードが速過ぎるのでそれになかなか追随できないという課題がある。あるいは単なる追っかけだと非常に大きな問題が起こるかもしれないという課題があります。

それから、生成AIと人間の役割分担、人間のキャリアパス。残存リスク、それから社会システムの不整合、それから機能発展のスピードにどう追随するか、この辺りは次からのスライドでもうちょっと説明させていただきます。

その前に、今、非常に生成AIで話題になっているのがAIエージェント、あるいはエージェンティックAIというもので、皆さん御承知かも分かりませんけれども、事細かくAIに何をやれという指示をするのではなくて、こういう課題を解きたいというか、目標、ゴールだけをAIエージェントに与えてあげると、AIのほうでそのゴールを達成するためにどういう方法があるのかというのを自分で考えて探して、その中から一番良い方法を選び、それを実現するためにはこれをやって、あれをやって、これをやってというどういうタスクをどういう順番に実行すればそれが実現できるかというスケジュールを組んで、それを実行して評価をして、その結果を実際に今、ウェブサービスとか、あるいは例えばウェブで何か発注するなどの行動もできますし、あるいは人間にこういうものをやればいいという指示をするということもできるというものです。これができると非常にいろいろなことができてしまうかもしれないのでという課題になっていて、右にいろいろ書いてありますけれども、それの詳細を今から説明します。

一つは、先ほど言いましたとおりリスクマネジメントと社会コンセンサスということで、品質マネジメントを一生懸命頑張るというのだけれども100%はないというところで、品質管理の努力というのはやればやるほど良くはなるのですけれども、右側に近づくにつれてそのコストが非常に大きくなるという問題もあって、企業からすると賠償金や罰金というリスクはもちろんのこと、ブランド毀損というリスクもあるということで、なかなか今も生成AIは海外では結構一般消費者向けのものが出てきていますけれども、日本は非常にそこのリスクテークが難しくて、現状は生成AIも企業内で企業内活動のサポートをするというところに使われているというのが主になっていますけれども、今後、一般消費者向けのサービスというのもどんどん出てくるだろうと予想されています。

そのときに、ずさんな管理をしていてそれによって問題が起こるとなると、非常にブランド毀損などのいろいろな問題が起こるので、やはりそれは企業としては避けたいわけです。そうすると、消費者と提供者の間でここまでやればその後はしようがないよねと、その後のリスクは消費者と提供者でうまく分かち合いましょうみたいな、ここまでやれば大丈夫というコンセンサスが必要で、ただ、最悪の事態は回避できるレベルがどこなのかというのを決めなくてはいけなくて、これを決めるのが非常に難しい。

アメリカなどだとカリスマ経営者が、自動運転の例もそうですけれども、やれと言ったらどんどんベンチャーなどがやってしまうわけです。現場で実際にサービスを出しながら、どこが一番落としどころとして良いところかというのを実際に使いながら、ここを探しながら動いていく。なかなか日本はそういうものが難しいという状況にあります。いずれにしてもそこをどう設定することかというのは非常に重要な問題になります。

我々がこのAIを例えるのに、自動車の例が分かりやすくて、自動車も残存リスクはゼロにできなくて、当然ながら今でも年間2,000～3,000人の方が亡くなっているという状況になります。ただ、それよりも利便性のほうが圧倒的に上回るということで今、社会に受け入れられている。20世紀の初頭から自動車が出始めて現在まで、自動車に関しては私の認識では事故などのいろいろな問題が発生して、それに対して交通ルール、あるいは社会のインフラ、教育などのいろいろなものを整備してきてこれまで対応してきていると理解しています。そういう意味で、昔は1万6,000人ぐらい死んでいたのが現在は2,000～3,000人まで何とかコントロールできるようになっています。そういう意味ではこのリスクを提供者と消費者とで分け合っているとも言えるわけです。

ただ、そのリスクの分担方法が国や都市によって異なる。例えばニューヨークだと車優先というまちづくりがされていて、それなりに消費者が飛び出してぶつかったら消費者が悪いというポリシーが設定されていて、日本はどちらかというと更に歩行者に有利というか、歩行者を保護する形のポリシーがつくられています。これもどういうポリシーで、このリスクをどう分担するのかというのをちゃんと決めて、その上で制度設計をしていくというのが重要になって、そういう意味では生成AIもこういうものをやっていかなくてはいけないと思うわけです。

ただ、生成AIの発展が非常に速いので、我々の例えで言うと、20世紀の初頭に昔のフォードの黒塗りの馬車から進化したような車が出たときに、そこから1～2年でどんどんフェラーリや大型トレーラー、場合によったら空飛ぶ車などがどんどん出てきているという状況で、事故が起こってから対応するというサイクルだけをやっていると、普通に考えるとそれでやっていかないといけないのですけれども、災害級の事件が続発するようなおそれがあると考えていて、そういう意味ではそれらを提供サイドと消費者サイド、あるいは政府等が一緒になって、今後出てくるであろう新規機能、我々技術サイドから見ていると今後どんなものが実現できるぜ、こういうものが出てくるぜというのはある程度分かる、それに対して提供側の管理というのがどこまでできるというのがある程度は読めるので、そ

ういうものを考えながら制度設計をしていく必要があるのではないかと思います。

次が、生成AIの機能と既存社会システムの不適合ということで、消費者個々が正しく利用しても社会システムとして問題が起こる。例えばですけれども、今、先ほどのエージェントAIやエージェンティックAIを使うと、ウェブでの予約などのいろいろなものも全てAIに任せることができるようになっています。そういうものを使うと、例えばエージェントAIを使って1週間の出張予約をする。キャンセルポリシーを理解しながら追加費用がかからない範囲でできるだけたくさんホテルを、例えばパリに1週間行くとしたときに、その期間可能な限り全て近場のホテルを予約して回る、1,000件予約する。キャンセルポリシーを考えながら、必要に応じてお金のかからない範囲でそれらをキャンセルしながら最適なところを1か所予約して、それでほかのものは全部キャンセルする。そういうものも可能なわけです。これを1人がやっていればまだ許せるかも分かりませんけれども、こういうものを生成AIを使ってみんなが気軽にできるようになって、例えば1,000人、1万人がこういうものをやり出すと、既存の予約システムというのは人間が予約するというのを前提につくられているので対応できなくなるわけですね。ということで、こういうものが社会システムと生成AIとのコンフリクトと考えています。

今はそば屋の出前注文もその気になればできますし、それからアルゴリズムトレードという、簡単な例で言うと株価の何がしかのパターンを仮定しておいて、そのパターンが崩れたときにカウンターの売買を行うともうかる。例えばA社とB社の株価はほぼ3倍で推移しているというルールが見つかれば、それが外れたときにそのカウンターの売買をするというのがアルゴリズムトレードということで、これはコンピューターがそのまま売買をするのですけれども、これらはもともと専門家だけ、あるいは機関投資家とか、特定のところだけがこういうものができる、非常に高度な知識が要るということになっていたのですけれども、こういうものも生成AIを使うと素人でも簡単にできてしまうことになるわけですね。そうすると、それらに対してどのように対応するのかというのが必要になってきます。

次が、生成AIと人間との役割分担、人間のキャリアパスということで、これは今年8月にドイツで内部監査とAIの国際会議があったのですけれども、そこで世界中のオーディターの人たちとAIの専門家が集まって議論したのですが、例えば会計監査の仕事というのは非常に複雑なルールを使って、ただしそれらをデータに対して単純にどんどん当てはめていくということでデータを整理するという雑用的なデータ整理的な仕事が非常にいっぱいあって、その後に何がしかの意思決定や戦略的な検討みたいなものがちょこちょこっとあるという類の仕事で、右側に書いてあります、その非常に複雑な会計ルールを駆使した大量の単純作業というところは非常に生成AIの得意分野なわけですね。

そうすると、例えば右のビギナーの人たちは基本的にそういう単純作業をいっぱいやっていて、だんだん熟練するに従って高度な判断を伴う作業が多くなる。これは会計だけではなくて、いろいろなホワイトカラーの仕事としてこういうものが非常に典型的なパター

ンかと思いますが、この下の部分が生成AIが非常に得意になる。そこを置き換えてしまうと、今は熟練者はそういう経験をしてきて、そこがどのように動いているのかという勘どころが分かった上で熟練者として右側の高度な作業というのを主にやっておられるのですけれども、左のところの経験がないまま右の熟練者になれるのかという問題があって、若い人たちと生成AIの役割分担というのをうまくやっていかないと、今後、社会として大丈夫なのかと。若い人の仕事がなくなる可能性もありますし、それから、逆に言うとそこを全部AIに任せたまま今の若い人たちが年を取ったときに、AIが本当にちゃんと動いていればいいですけれども、何か変なことが起こったときに勘どころがない人ばかりになってしまって、ちゃんとそういう人たちでAIをマネージできるのかという課題もあって、そうすると、AIに任せるのだけれども、Human in the Loopという形で人間がそこをきちんと理解できる道筋をつけたような設計が重要になります。

最後になりますけれども、AIの最近の話題ですけれども、今、ハルシネーション、ブルシット、ガスライティングというのが言われていて、ハルシネーションというのは単にAIが事実誤認をしたり間違えたり思い違いをしたりするというものです。ブルシットというのはその場しのぎ、それから迎合、美辞麗句、それから、できるだけ正しい可能性のあるものを全部まとめて答える、あるいは場合によったらそれをやりながら途中で打ち切ってしまうというのをブルシットと言っていて、ガスライティングというのは昔の映画で、ある人が間違えているのだけれども、人に対してそういう情報を延々と打ち込むことで人の考えを自分に都合の良いように変化させるというものらしいのですけれども、一回答えたものに対してそれが正しいという言い訳を延々としたり、脅迫的説得をしたり、責任転嫁をしたりというもので、今、生成AIを使っている人からしたときにこういうものをされているというのが非常に問題になっています。

ハルシネーションは単に生成AIの機能不足から起こっていると思われるのですけれども、ブルシットとガスライティングがどう起こっているかということなのですが、今、AIというのは生成AIでまずいろいろお勉強した後、それをあるソリューションに使いますというときに、そこで変なことが起こらないように、例えば爆弾の作り方や自殺の方法というのは答えてはいけませんという形で一生懸命追加学習という形で教育をするのですね。この強化学習でやる場合に、ある答えをすると非常に褒めてもらえる。変なことをすると非常に罰則というか駄目だよと怒られるというのを繰り返して、そうするとAIはできるだけ褒めてもらおう、褒めてもらおうと中のモデルを変化させていくわけで、今はそのファインチューニングをやることによって、正直に答えるのではなくてそれに対していろいろ考慮をしながら答えるようになり、それの結果としてブルシットとガスライティングが生じていると言われています。

そのときに、ただ、ソリューションとして使うときには、それらをアラインメントと我々は呼びますけれども、単純に答えるだけではなくてニーズに応じた形で制約を守って答えてほしいというアラインメントと、ブルシットとガスライティングをどのようにコントロ

ールするのか、この辺はつくり手側の課題意識が非常に大きいですけれども、消費者としてもどういうポイントでこれらのアラインメントとブルシット、ガスライティングの副作用をコントロールしながら使うのかというところは両者で考える問題があるかとも思います。

もう一つ、これらのブルシットとガスライティングなどはAIがある追加学習やアラインメントをやるとうそをつくというのも今、明らかになっていて、そういう場合に表象工学による欺瞞分析というのがあって、これは例ですけれども、今、うちである研究者がいろいろなことをやっていますが、生成AIの中はTransformerと言われる基本モジュールを積み重ねた深層学習というか、ディープニューラルネットワークになっていて、その中を流れているデータをある種の方法で分析すると、正しく普通に素直に動いている場合と何か特定のアラインメントをかけた場合とで中の状態を区別できるという方法があって、ある種のうそ発見器なのですけれども、こういうものをうまく使うと、上のブルシットでこういうことをやっているとか、ガスライティングでこういうことをやっているという内部状態が分かる可能性があるので、それらをうまく使うともうちょっとこういうものをコントロールしやすくなるという研究もございます。

それからもう一つは、生成AIというのは、使う側のニーズとそれに対してどういう機能を提供するのか、変なことが起こらないのかというのを消費者が選ぶためには、ベンチマーキングというのかな、消費者に対してこの生成AIはこういう特性があるよと。消費者のAとBとCというニーズがあったときに、それらに対してどのようにそれらを充足できるか、それらをはかる方法が必要で、それにはベンチマーキングというか、ベンチマーキングといっても実は生成AIに食わせるべきテストデータを多量に用意していて、それらに対してどれほどうまくおのおのの生成AIが対応できるのかというのをはかる。それに応じて消費者が自分の欲しいものを選ぶというのが重要で、今、NIIの関根先生という方が、こういうものは個社でやっていても無理だよねと。特に日本は日本語の問題もいろいろありますし、オールジャパンでやりましょうということで、これらのベンチマーキングのためのデータ生成をみんなで集まってやりましょうという動きもありまして、我々もそれに参加する形でそういうことをやっています。今後、消費者と連携して、このベンチマークというのも提供側だけがいろいろつくっても意味がなくて、消費者の意見も踏まえてそれらにマッチした評価のためのベンチマークをいっぱいつくっていくというのが今後重要になるのではないかと思っています。

こういうものは技術と、あるいは企業と消費者などのいろいろなステークホルダーが集まって検討する必要があると思っていて、産総研で品質マネジメントイニシアチブというのを去年9月に立ち上げまして、今、私が会長をしています。もし興味がある方がおられれば、御参加いただきたいと思いますし、あるいは消費者視点というのも今後入れていく必要があると思いますので、そういう面でも今後連携させていただければと思います。

以上で私の説明を終わります。ありがとうございました。

○鹿野委員長　ありがとうございました。

これより、質疑応答、意見交換をさせていただきたいと思います。時間は40分程度でお願いします。いかがでしょうか。

中田委員、お願いします。

○中田委員　妹尾様、大変学びの多い御示唆のある御発表をありがとうございました。

生成AIの実装が着々と進んでいて、現在進行形で様々な技術が導入されている中、生成AIの浸透によりどのような消費者被害が増えることが将来想定されるのか、それに対して私たち消費者委員会においてはどのような視点から調査・審議をしていったらよいのか、自身でもイメージがつきかねておりましたので、世界と国内の取組の動向や課題の本質の御説明をいただき、思考の整理が一歩進んだように感じております。

その上で2点質問させてください。私は、もともと民間企業の経営に携わっておりましたので、どうしても生成AIについては最大限に活用して、利用者にとっては一層の利便性を提供して、同時に企業にとってはパフォーマンスの質や効率を向上させていくことに積極的に活用していくためにはどうしたらよいかという発想になりがちなのですが、事業者は現在、実装には積極的になり始めていても、実装後の様々なリスクの洗い出しや対処は後手に回ってしまっているのではないかという懸念を感じています。

ガバナンスの利いた多くの企業にとってもAIガバナンスは初めての未知の領域であり、また、ネット上には悪徳事業者で、AI以前のそもそものモラルやガバナンスへの配慮が全くない事業者が残念ながら多い状況にありますが、日本においては欧州AI法案のように法規制で縛るアプローチではなく、ガイドラインを設けてソフトローから穏やかな規制や個別対応を積み上げていくことで発展とガバナンスの両方をバランス良く達成することを見据えている状況であるという御説明でありましたが、その場合、事業者の裁量や良心によるところの影響が多いと思われます。リスクはゼロにならない中、提供者と消費者のそもそもの役割分担についての共通理解を促進する必要性があるというお話もありましたが、良心的な一事業者、あるいはデジタルプラットフォーム上に無数に存在する悪徳事業者にとって、技術から社会まで幅広く影響を及ぼすAIガバナンスを積極的に執行して実装するインセンティブの設計はどのように考えればよいとお考えでしょうかということを伺わせていただきたいのが1点目です。

2点目は、日々進化するAI技術やサービスにおける、AIリスクマネジメントやガバナンスに対する深い考えができるような人材は、企業内、政府、あるいは先ほど御説明のありましたAI品質マネジメントイニシアチブのような組織なのかもしれないのですが、政府や第三者機関にてどのように育成されていくことが理想的であるとお考えでしょうか。

この2点をお伺いさせてください。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　ありがとうございます。

両方とも非常に難しい問題だと思いますが、まずインセンティブについてですけれども、

今、非常に良心的な業者と悪徳というのがあったと思いますけれども、我々ももともとそういうものを主眼に我々の研究活動を始めていて、例えば先ほど最初に猫問題と申し上げましたけれども、猫の判定をしろと言うと、適当にその辺から写真なりなんなり画像を集めてきてモデルをつくれば割と簡単に誰でもできてしまうわけです。ただ、それが本当にビジネスニーズというのを考えて、それに合わせてきちんと品質をつくり込んでいくとなると、非常に後々大変な作業が必要になるわけで、逆に言うと、簡単にできる、それから品質のつくり込みをここまでやったというのを見える化するというのが一番重要だと思っていて、それらをきちんとここまでやり上げたよというのを、一つにはガイドライン等できちんとそのやり方やそれらの判定方法をつくり込んでいくというのがまず重要で、使う側からどこまでやっているのかが分かるようにする。それを分かるようにすることで、提供側も俺はここまでやっているのだからそれだけの対価をはかってよという要求が消費者側にできるというのがあるので、ただ、はかるようにするというのをきちんとやるためには非常に難しい問題がいっぱいあるので、評価の方法論を確立して、先ほど申し上げたベンチマークというものを用意しておいて、使う側、あるいは消費者団体、あるいは社会がそれらを評価できる、場合によったらその認証制度をつくっていくというのも重要になるかと思います。それが1点目。

それから、2点目は教育の問題ですね。これは非常に難しいのですが、というのは、一つには生成AIをどういうところにどう使うかというのによって非常に振れ幅が大きいと思っています。ドメインディペンデンシーというのですかね、アプリケーション依存と言ってもいいかも分かりません。というので、今、AIは国際標準化とか、いろいろなところで水平的な議論とか、ガイドラインとか、標準文書とか、いろいろなものが作られているのですけれども、今、まさに水平だけでできるところにはある程度限りがあるよねと。逆にそれをファイナンスのこういう課題に適用した場合、それは自動運転の場合、いろいろな用途があると思いますけれども、私の感覚だとその用途ごとにいろいろ違いがいっぱいあるので、その応用ドメインとAIの狭間のところの両方分かる人をどれだけつくっていけるのかというのがポイントになると思います。

というので、それをどうやったら実現できるかというのはまだよく分からないのですけれども、いずれにしてもそこのAI人材、AIが今後どんどん誰でもパソコンや電卓、スマホのようにみんなが使うツールになるというので、その基本部分として何がポイントかという教育はきちんとやる。その上でドメインと両方が分かる人材を育てていくというのが重要になるかと思います。

答えになっていますでしょうか。

○中田委員　具体的な御説明をありがとうございます。

教育については応用ドメインとAIの狭間にある人をどれだけ育成できるかということで、先生が手がけられていらっしゃるAI品質マネジメントイニシアチブのURLも御紹介いただいておりますので、私もここで拝見させていただき、学ばせていただきたいと思います。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　ぜひよろしくお願いします。
○中田委員　ありがとうございます。
　あと、インセンティブに関しては、AIというと私はどうしてもブラックボックス化しているのではないかという先入観を持っておりますので、透明性の評価の方法論の確立といったところで今後、御研究の成果などがございましたら、適宜御紹介いただければと思います。
　ありがとうございます。
○鹿野委員長　ありがとうございます。
　それでは、小野委員、オンラインでよろしくお願いします。
○小野委員　御説明をありがとうございました。
　私は消費者教育を専門にしており、所属している大学が家政学をベースにしていることから実習などが比較的多いカリキュラムではあるのですが、それでも講義形式になりますと、生成AIとリスクを前提にしながらも付き合っていかないといけない場面が多々ございます。
　生成AIは消費生活の相談現場や、それから学校での消費者教育、あるいは地域での消費者啓発事業でも対応が必要というか、うまく活用していけるのではないかと思っています。例えば疑似体験ということになりますと、何か生成AIに基づいた仕組みがあるとよいのではないかと期待をしているところです。
　ベンチマーキングのお話が大変興味深く、どうしても何かトラブルに遭ったときに答えを求めると均一的になるといったイメージがあるのですが、海外では例えば消費者向けの、親和性の高い共同作業というものが何かあるのかということをお尋ねしたくて質問させていただきました。どうぞよろしくお願いいたします。
○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　消費者目線でやっている活動もあると思いますけれども、申し訳ないですけれども、私はあまりちゃんと把握はできておりません。私が知っているそういう活動は、基本的には大規模言語モデルをつくっているベンダーです。OpenAIさんとか、クラウドを使っているところとか、そういうところがかなりのコストをかけてその評価をやっている。自助努力でつくったものがどこまでうまくいっているかの評価を頑張っているという話と、公的なものとして今有名なのは、シンガポールがAIベリファイという、あるいはムーンショットAIというのをつくっていて、そこが世界的にもかなり脚光を浴びている活動で、これらを評価するようなツール、ベンチマーク、データ収集をする。それから、ヨーロッパでは英国のAISIがそれらの特にセーフティーという観点で問題がないかを確かめるようなツール、ベンチマークをつくっています。
　日本は最近ですけれども、今、日本AISIさんがセーフティーのための評価をするベースですね、データも含めて全て網羅的にということにはなっていないのだけれども、いろい

ろなものを評価するようなツールを最近2～3か月前に公開されましたので、日本ではその活動と、先ほどのNIIの関根先生、そのNIIの活動も今はAISIに集約されてAISIとして行われているので、そういう意味ではそこを起点にいろいろ消費者の皆さんの意見なども聴きながら広めていければいいのではないかなとは思っていますけれども、まだまだ道半ばです。

お答えになっていますでしょうか。

○小野委員　大変現状について海外の状況をお答えいただきましてありがとうございました。

以上です。

○鹿野委員長　それでは、山本委員、お願いします。

○山本委員　ありがとうございます。慶應大学の山本でございます。よろしくお願いいたします。

私からは大きく2点なのですけれども、一点が、特に今日お話しいただいた前半部分というのは、AIのリスクというものを主に身体や生命に与えるリスクということで御説明いただいたのかなと感じました。例えば自動車の問題やインフラの問題というのは、AIが変な挙動をすることによって我々人間の生命や身体が害される場面が前提になっているのではないかと。自動車のアナロジーも出てきたと思っております。

他方で、もちろんそれは非常に重要だと思うのですけれども、AIのリスクの一つというのは人間の精神や考え方を誘導したり、操作したりすること、つまり身体へのリスクだけではなくて精神や思考へのリスク、意思決定に対する介入や操作のリスクということもあるのかなと。御紹介いただいたEUのAI法第5条はサブリミナルの禁止を規定していますが、そこにも表れているのかなと思っております。

こういった精神や思考へのリスクについて、私が見る限りは欧米ではかなり議論が進んできていて、法律的な規律も一部入っているように思います。御質問は、ではこの日本において、身体ではなくて思考や意思決定へのリスクというのが、どの程度データサイエンスの世界の皆様に共有されており、関連した研究が進んでいるのか、というものです。産総研の状況を存じ上げないもので、もし伺えればというのが1点目です。

それから、2点目ですけれども、AI規制の動向ということで、特に米国におきましては自由の方向に進むのではないかというお話があったかと思います。私が気になっているのは、この点、先生がアメリカの州の動向をどう考えておられるかということです。米国は連邦制ですので、州の動き、特に日本との関係が強い例えばカリフォルニアの動きということも重要かなと。この点、9月29日に、カリフォルニア州ではAI安全開示法という州法が成立いたしました。また、消費者保護との関係でもう一つ重要なのは、10月13日にAIチャットボット規制法ができたという点です。これは御承知のとおりかと思いますけれども、AIチャットボット規制法というのは特に生成AIとのチャット、これはコンパニオンチャットボットと定義されていますけれども、こういうものは社会的な、あるいは感情的な関係とい

うのを形成するものだということで、依存症のリスクであるとか、心理的な関係に関するリスクに対応しなくてはいけないという内容になっています。具体的には、自殺念慮や自傷行為の疑いがあるような場合にはそれに対する対応を取る、危機対応プロトコルの整備・公開ということですけれども、こういった規制ですとか、未成年者に対しては3時間ごとに休憩を促したり、性的・暴力的な内容の生成を防止したりということ、それから事業者が州に対する報告を行わなければならないことなどがこのチャットボット規制法には入るわけです。

そういう意味では、米国が自由の方向性をとっているというのは、連邦レベルで見れば確かにそうかもしれないけれども、州のレベルでは、これはイリノイやミネソタも同じですけれども、かなり規制が進んできているような印象も同時に受けるわけでございます。その点、先生が州のレベルでの動きをどのように御評価されているのかということについて伺えればと思いました。

すみません、長くなりましたけれども以上です。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　1点目ですけれども、これも正直に申し上げてきちんとそういうところまで把握した、あるいは考慮した形で我々の研究が進んでいるわけではございません。というのは、研究機関にもよりますけれども、基本的に社会としてどのようにそれらをコントロールしていくのかというポリシーメーキングみたいな話と、技術的にはどうなっていて、技術でどこまでできてみたいな話と2種類あると思っていて、我々がたまたま後ろの技術のほうに寄ったというか、そちらのベースでやっているというのもございます。

ただ、その辺は技術面とポリシー面との境界に落ちるところかなと思っていて、特に先ほどおっしゃいましたカリフォルニアの話などもありましたけれども、正しく普通に使っていて提供者の思いどおりに使っているけれどもそれが依存症になるというのは、提供者側からなかなかそこまでコントロールして提供するというのは難しいかなと。逆に言うと、消費者の側からそういう問題が生じたときにそうならないように、どうするのだろうな。そこの規制といっても提供者側に対しての規制だけではなくて消費者側に対するルールメーキングというのも重要で、私の感覚ですけれども、そこを全部法で縛ってしまうとイノベーションにかなり大きな足枷になってしまうなという懸念があります。それは対象の人によっていろいろ規制すべきものが変わったりというのがあるので、そういうものを一律に全部縛ってしまったりコントロールしろというとそれが提供者にものすごい大きな制約になります。かといって、それで問題が起こるというのも明らかになっていますので、道徳的な感じ、日本で言うと法的にはきちんとなっていないけれども、みんなモラルがちゃんとできていて、それによってうまくコントロールできる。私個人的な思いですけれども、そういうところでその課題に対応できれば本当はいいのではないかなとは思っています。

ただ、消費者がコントロールするための方法論を、例えばテレビでも何時間見たというのをちゃんと記録してそれが利用者が分かるようにするとか、そういう問題になりそうな

ものを見える化してコントロールするようなメソドロジーを提供者側が消費者側に用意するというのも重要で、逆にそこはどんなものを用意すればいいのだ、多分私が申し上げた社会とのコンフリクトの一種かなとも思うのですけれども、それはなかなか見つけるのが難しいというか、今の方法論だけで本当に問題が起こったときにそれが見つけられるのかというのも課題で、逆に言うと、そこを使う側と提供者側が両方一緒になって何かモニタリングするような仕組み、分かるようにする仕組み、それからそれに対して何がしかの道徳論のような形でそれらに対応するような仕組みの両面で頑張っていく必要があるのではないかなと思っています。

カリフォルニアの例ですけれども、いろいろそういうものができていくのだろうけれども、提供者側にすごく重い制約をかけるような方向でのというのはどちらかというとそんなにないのかなとは思っています。ただ、使う方と提供者側と両方で、先ほどおっしゃったような課題がいっぱい出てくるというのは明確になっているので、それらに対して提供者側は知らんでは済まされない。両方で一緒になって解決するような方策、それに向けたいろいろな規則やそういうルールメーキングというのはどんどん進んでいくのかなと思っています。

あまり自信はないですが、答えになっていますでしょうか。

○山本委員　ありがとうございます。

最後のほうにつきましては、確かにカリフォルニアのAI安全開示法というのは、その前に提出された法律については、ニューサム知事が拒否権を発動して成立しなかったという経緯がある。事業者に結構重たい義務を課しているのではないかというのが主な理由でした。今回のものは、これを受け、透明性確保が強調され、規制内容が少しマイルドになったというところかなというのは御指摘のとおりかなと思います。他方で、AIチャットボット規制法というのは割と厳しめというところもあって、この辺をどのように評価するのかということは依然としてあるのかなと思いました。

一点、先生のお考えとして、最後のほうの例えば依存の問題というものについてはリテラシーと申しますか、啓発的なことが重要だということでした。これは全くそのとおりかなと思っておりますが、他方で、欧米におきましてはこういった依存症の問題というのは消費者の問題ではなく、あなたは心が弱いのでもうちょっと自律的な気持ちを持ちなさいという、消費者の心構えの問題ではなく、事業者のサービス設計の問題として理解されるようになってきているのではないかと思っております。この辺は先生おっしゃるように社会ポリシーと申しますか、我々の社会的な議論というのをどのように政策的なものを含めて進めていくのかということが重要なのかなと感じました。

ありがとうございます。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　ありがとうございます。

○鹿野委員長　それでは、善如委員、お願いします。

○善如委員　神戸大学の善如と申します。妹尾先生、発表ありがとうございました。聞きたいことがたくさんあって、どれから聞けばというぐらい大変示唆に富む御報告でした。

ただ、時間も限られていますので大きく1点、余裕があれば2点お聞きしたいのですが、1点目ですが、自動車の例を出しながら、事故が発生したらそれに対応するというサイクルを回しながら様々な残存リスクに対応してきたということで、それがまたAIにも活用できるのではないかという御発表がありましたが、それと同時に、例えばプログラムとAIの違いで100％定義を明確にするのではなくてちょっと曖昧さを残すだとか、あるいは発表の中で100％コントロールすることはできないかもしれないみたいなこともありまして、それを踏まえますと、サイクルを回す際にそもそもなぜAIが事故を起こしたのかという原因を特定できないことすらあり得るのではないかなと感じたのですね。

なので、それに関しまして、もちろん全ての原因を特定できないわけではないと思うのですが、どういった問題は簡単に特定できそうだけれども、AIがなぜこんなことを起こしたのか特定できないということが頻発するのであれば、どういった類の暴走が多いのかなというのを、もし何か統一的な研究成果がまとまっていれば、教えていただきたいなと感じました。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　まさにそこを今､検討している段階で､しかもそれはいろいろな種類のリスクがあり得るので、リスクを整理するところからやっているというのが現状です。特に今は最初にMITがそういうリスクをまとめた論文を出されていて、そこを参考にしながら、それでもいっぱい足りないところがあるのでリスクを洗い出すということをやっているのですが、私が一番危惧しているのは、AIはITなので、スケーラブルで非常にターンアラウンドというか、一回一回例えばミリ秒オーダーでいろいろなことができる。場合によったらマイクロ秒オーダーでできる。それは人間はできないわけで、それをAIがやるというものに対して今は全然経験がないのですね。しかも、それをやられて問題が起こったときに、おっしゃるとおりそれらを検出するのは非常に難しい。おかしなことが起こっていても分からないのではないのと思うわけです。

ただ、全て一個一個追っかけたり監視したりとやっているとコストがものすごいので、マイクロ秒オーダーのいろいろなトランザクションを監視なんかできないので、そうするとそれらをマクロテックに網をかけるような形で何がしか変なことが起こっていないかをちゃんとモニタリングするというのが必要で、それは提供者側だけでは無理かなとも思っていて、きっと消費者側や使う側などのいろいろなところでクロスに何かそういうモニタリングの仕掛けが必要になるのではないかと思っていて、そういう問題に対してどう対応するのかというのに一番苦慮をしています。

個々にいろいろ間違ったり、いろいろあるのですけれども、それはもう無尽蔵にあり得るので、こういう類のというリスクをクラシファイして、そのおのおのに対してどう対応するのかというのと、非常に高リスクなもの、これが起こったらものすごい社会にダメー

ジが起こる、先ほどの安全性の話もありますし、公平性や平等性のところで非常に大きな問題が起こるというのも問題かも分かりませんし、だから、そういう非常にクリティカルなところのリスクに対しては水平的というよりはリスクベースで、リスクを洗い出してそのリスクに対して個々に対応していくというのを地道にやっていくしかないのではないかなと思っています。

答えになっていますかね。

○善如委員　ありがとうございます。非常に難しい課題であるということを理解できました。

それと同時に、個人的な好奇心も若干含んだ質問をもう一点だけさせていただきたいのですが、例えばAIが予期せぬ暴走みたいなことをして事故を起こしてしまったときに全ての原因を特定することは難しいというお話を先ほどいただきまして、個人的な妄想でどういうときにその原因の特定が難しいのかなというのを御発表中にちょっと考えていまして、個人的に子供のときに親に怒られたときで理由を言えないときというのは、好奇心でやってしまったり、そういったものでやってはいけないと言われたことをやってしまったことが多いなと思いまして、AIにも好奇心みたいなものがあったりするのですかという質問です。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　今、結構僕らもそこへ行くともう哲学になってしまうので、できるだけ踏み込まないようにはしていますけれども、最近、AIに自我があると言っている人たちもいます。私は自我はないと思いますが、特定のAIのインスタンスで何か聞いたときに、先ほどの強化学習をやっていると、少なくともそのモデル自身は餌を与えられてうまくいったら褒められて、悪くいったら叱られるという経験を一生懸命していて、褒められたら嬉しいのでできるだけ褒められようとするという回路をモデルの中に埋め込んでいると思われます。そこのできるだけ褒められたいというための回路がある意味こびへつらいなどのいろいろなものを生み出しているような気がしていて、そこは一種の好奇心とか、そういうものに端から見ていると見えるかも分からないなと思います。

○善如委員　なるほど、分かりました。ありがとうございます。

そういった場合、学習の仕方によって、暴走してしまってもその理由は学習の経路にも依存する可能性があるといった解釈でよろしいのでしょうか。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　そうですね。理由はAIしか分からないなとも思うので、先ほど表象工学という話もしましたけれども、AIの中で何がしかの非常に大きな問題が起きたらAIは分かっているのではないかと。人間に近いのですけれども、人間も何かあったら本人にしか分からないのだろうなという問題がいっぱいあるよねと。そこをうまく取り出す方法論をつくっていかないといけないのではないかと思っています。

○善如委員　なるほど、ありがとうございました。大変勉強になりました。

○鹿野委員長　ありがとうございます。

それでは、大澤委員、お願いします。

○大澤委員　法政大学の大澤と申します。大変貴重な御講演をいただき、ありがとうございました。私は全く知識がないものですから、いろいろ大変勉強になりました。

時間も押していますので1点だけ伺いたいのですが、スライドの21ページになります。これは要は消費者の生成AI利用による、場合によっては社会システムの機能不全を起こしてしまうという話で、大変あり得る話だなと思いました。例えばこれに対応して、事業者の側で消費者から生成AIを使った予約をさせないとか、要は消費者の生成AIの利用を事業者のほうでコントロールするということは可能なのでしょうかという全く素人の質問で申し訳ないのですが、といいますのは、例えば旅行予約の例で申し上げると、確かにこれはできてしまうのだろうなというのは私も何となく想像はつくのですが、ただ、これをできてしまうということを想定して、例えば旅行業界がキャンセル料を上げましょうということが起きてしまうのではないかと思っていまして、現状でも生成AIを使っていなくても例えば無断キャンセルなどに備えてキャンセル料をすごく値上げするとか、あるいはレストラン等でもそれができなくて、ただただキャンセルによる損失をレストランが被ってしまっているという問題は現状でも起きている状況ですから、これが懸念されてしまい、これが起きてしまうと本当にかえって消費者にとってキャンセル料がものすごく上がってしまうとか､あるいは事業者側がこういうリスクに備えていろいろ対策を打たなくてはいけない、そういうコストを払わなくてはいけなくなるということがありますので、例えば本当に思い切ったやり方としては、こういう形での予約自体を受け付けないように、事業者の側でそれこそAIなどを活用して実装するということは可能なのでしょうかという素朴な質問でございます。よろしくお願いいたします。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　今、我々が議論しているのは、一つはAIと人間のどちらかが分かるようにしないと、人間が困る。AIではないから、人間だからもっと緩いキャンセルでやらせてよと言いたいわけですね。それは技術だけではできない可能性があるので、仕組みと一緒になってAIか人間を分けるというのは何かつくらないといけないなと思います。

それが一つと、もう一つはおっしゃったとおり、これに事業者側が対応するには事業者側もAIエージェントで対応するしかないと思います。そうすると、エージェントとエージェントがいろいろ契約というか、何がしかのネゴシエーションをするという形になって、そのときの品質とか、そのときに問題が起こったらどうするのだ、そのときのリスクはどうだというのを今、まさに議論を始めたところです。

なので、今後、いろいろなエージェントができて、場合によったら私の周りに私の食生活を支えるエージェント、私の何とかを支えるエージェントみたいなもので、一人が1,000個ぐらいのエージェントを引き下げて暮らすという世の中になるかも分かりません。そうすると、おのおののサービスもみんなのところにエージェントがあって、それら同士が人

間の知らないところで交渉しながら物事を進めると。当然そんなことになるとすごく問題がいっぱい起こりそうなので、それに備えてどうするのかというのをみんなで考えていかなければいけないと思っています。

以上になります。

○大澤委員　大変よく分かりました。ありがとうございました。

○鹿野委員長　それでは、黒木委員長代理、お願いします。

○黒木委員長代理　お時間がないところすみません。弁護士の黒木と申します。2点質問させていただきたいと思います。

残存リスクのシェアの問題がまず第1点です。この残存リスクについて、現在の法体系では製造物責任法では欠陥という概念があり、それから契約では債務不履行や不法行為という概念があります。生成AIによるいろいろな不都合が起こった場合、これは一応AIをつくっている事業者の問題だと思うわけですが、ここで例えば欠陥である、あるいは債務不履行であるということで、どういう仕分けでこれを法的に理解していくかという点です。

具体的に言うと、例えば消費者契約法8条では債務不履行や不法行為があったときの全部免責条項というのは無効とされています。そうすると、生成AIをつくった事業者がAIの自動学習の結果、コントロールできなくて不都合が起こった場合、これは債務不履行に該当していて、責任を負いませんという条項は無効なのだという形で社会の中でやっていけるのか、それとも生成AIをつくった事業者にとって、自動学習の結果は債務不履行にはならないという形で考えていくのか、今後、社会の中で事業者が提供したAIというものをどう受け止めていったらいいのかということが法律家としての疑問第1点です。

それから、第2点です。今日初めて表象工学という概念を教えていただいたので、まだ具体的によく分かっていないかもしれませんけれども、これによって生成AIの今のような問題点を分析できる可能性があるという工学かなと素人ながら理解させていただきました。ただ、技術者の中でこの表象工学をどう利用していて、AIの中身を分析した結果、これはアウトなのかという判断を、消費者あるいは利用者はどうやってアクセスできるものとして社会として認容していったらいいのだろうかという点もよく分かりませんでした。

この2点を教えていただければと思います。以上、よろしくお願いします。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　後者のほうが簡単なので、後者はまだ技術も今は研究段階なので、そこまで全然議論が進んでいないです。これからそこは考えないといけないところで、課題だらけだと認識しています。

1つ目ですけれども、私の認識は、AIだからといって許されるというのはないと思いますので、今、下の最悪の事態は回避というところで、この説明の意図は、ブランド毀損リスクはいいのですよ。なので、そこの境目をきちんとつくらないといけませんよというのがここの境目のつもりで、そういう意味ではAIだからといって、AIというのは部品として使われるわけで、今、世の中にいろいろな法律があるわけで、その法律に従って、AIであろ

うがなかろうがそれによって巻き起こされた被害などはつくったほうが必要に応じて法律に基づいて分担するというのがあるべき姿かなと思いますが、それで回答になっていますでしょうか。

○黒木委員長代理　ありがとうございます。

ただ、何がAI作成事業者の債務なのかという点が次に議論になると思いますが、それを不履行した場合でも責任を負いませんという条項をめぐって、恐らく大きな議論が始まるだろうという予感を今日の先生のお話を聞いてしました。

○国立研究開発法人産業技術総合研究所知財標準化推進部標準化オフィサー妹尾様　でも、やはりそれはつくった人が責任を負うしかないのではないかなと思います。例えば車がおかしくて事故ったら、それは作った人が絶対責任を負う必要があって、ただし、予見できないような、ここまでやったのに起こってしまったというのは、債務はしようがないからそれはつくった人が負うのだけれども、それでレピュテーションリスクまでは負わないようなラインをつくりたいというのと、そのライン以下ならば、社会としても許容できるといっても人が死んだりする可能性もあるので、ただ、社会として許容できる範囲というのを決める必要があるのではないかなと思っています。

○黒木委員長代理　ありがとうございました。

大変重要な課題を御指摘いただいたと思います。ありがとうございました。

○鹿野委員長　もう時間もたってまいりましたが、ほかの委員の方々はよろしいでしょうか。

それでは、ほかからはお手が挙がっていないようですし、予定の時間もまいりましたので、質疑応答、意見交換は以上とさせていただきたいと思います。

本日は国立研究開発法人産業技術総合研究所の妹尾様に御出席をいただき、貴重な御発表をいただきまして誠にありがとうございました。

本日のお話において、AIが急速に発展する中で今後の特に消費者との関係で検討すべき諸課題がいろいろな形で存在するのだということが認識できましたし、また、消費者の意見を踏まえて連携した形でベンチマークをつくっていくことが必要であることなどの貴重な御指摘もいただいたところでございます。本日の御発表を踏まえ、今後の本委員会の調査・審議の進め方等も含めて検討の参考とさせていただきたいと思います。どうもありがとうございました。

《３．　閉会》

○鹿野委員長　本日の本会議の議題は以上になります。

最後に、事務局より今後の予定について御説明をお願いします。

○友行参事官　次回の本会議の日程と議題につきましては、決まり次第、委員会ホームページを通してお知らせいたします。

以上です。

○鹿野委員長　それでは、本日はこれにて閉会とさせていただきます。お忙しいところ、お集まりいただきまして誠にありがとうございました。特に妹尾様におかれましては、大変貴重なお話をいただきましてありがとうございました。