第 2 回　AI 社会実装アーキテクチャー検討会　議事概要

2020 年 9 月 7 日　10:00-12:00

- **AI と金融規制における議論の類似性に関する論点**
  - 本ガイドの議論は、リスクベース、プリンシプルベース、アカウンタビリティなど、金融の規制における議論と類似性がある。ただし金融はリスク自体をやりとりするものである。AI の議論と金融の議論をどう整理するか。
  - 金融のガイドラインは、規制庁が国際的なガイドラインを参照して作成したものである。法的な世界では不法行為に関する議論が中心。その一方で、AI が実装されていく世界は契約ベースで、契約責任が中心になる。

- **リスクベース・法的責任の考え方に関する論点**
  - リスクベースのアプローチについて。対置する概念としては、ハザードベースアプローチというものがある。生じる害だけ見て規制をするのがハザードベース。毒をとにかく排除する考え方である。そうではなく、リスクを考えるべきとだというのが、リスクアプローチである。「AI はハザードベースだと排除すべきとなってしまうので、リスクベースと考えるべき」という説明が強調されてもよいのではないか。
  - 「リスク」と言う言葉の位置づけについて、今の文脈では多義的ではないだろうか。ヨーロッパは、ワンサイズフィッツオールではなく、リスクの高い AI に対しては、高いリソースをかけて取り組んでいく、逆もしかりで、全体で強度を決めていくという考えである。
  - 法的責任がどうなるかは重要な論点。ソフトローのインセンティブとしてポジティブに位置づけられることもある。EU では、プラットフォーム規制の文脈で、具体的にはあれこれの ISO 基準を参照して決めるべしという記述がある。
  - ガイドラインにおいて、法的責任の話をする際、技術的に何ができて何ができないという前提なのかをクリアにしたい。暗黙的にできると認識されていて、あとになってできなかった、という話になると良くない。
  - AI の利活用を促進するうえで、合理的に過度の責任を負わせない仕組みを設けるという点は賛同する。一方で AI はブラックボックスであり、利用者側にとって信頼性やリスクを評価することは難しい。どのように監査や外部機関の評価をするかを示していけるかが重要である。実施体制そのものについては言及されているが、ユーザーにとっては、評価するのは企業であっても、プロジェクトごとに評価する仕組みを考えるのは難しいものだろうか。
  - ガイドラインに、AI と一般的なソフトウェアの運用ルールが異なるという点をより強調してもよい。実務では、AI と旧来型ソフトウェアを同じルールに乗せてしまうことでうまくいかないことが頻発している。
  - ガイドではリスクベースの考え方など納得する部分がある。一方で、いざこれを実施しようとした際、社会実装までは差があるのかと思う。仮に自社で取り組むとすれば、製品分野ごとにリスクがどの程度あるのかを考え、ここまでやるようにと指示することになる。可能であれば、分野ごとに、例えば自動運転はこうだとか、具体的な例まで詳細化できるとよい。

- 「監査」を含むアシュアランスの方法はリスクや役割・責任範囲などに応じて、自己チェック、ピアレビュー、内部監査、第三者による監査など様々な様態を想定しておくべきではないか。リスクや影響が小さいと合理的に評価できる場合、必ずしも「監査」まで求める必要がないケースが想定されるため。また、AI を用いたサービスやプロダクトの開発・運用には 中小企業やスタートアップを含む多くのステークホルダーがおり、それぞれの役割・責任範囲は異なることが想定されるため。

- **多様な事業者・事業分野等が存在するなかでのリスクの考え方に関する論点**
  - 本ガイドラインが対象にするのは全体であると認識しているが、大企業、中企業、小企業、スタートアップでは、それぞれ置かれている環境が違う。METI のプライバシーガバナンスのガイドでは、そこは柔軟に様々な企業の存在を前提に考えていた。
  - 事業分野ごとにリスクの性質が違ってくる。リスクの性質に合わせて、求められるリテラシーや体制の整備は違ってくる。こういうリスクが予想される場合には、こういった体制が必要という、なにかしら示すものがあると、読み手となる企業担当者は理解がしやすい。

- **複数間企業において AI システムを開発する際の論点**
  - 大企業を除いて、AI システムは 1 社で完結せずに複数の事業者で成立する。特にデータのラベリングなどでは、AI に関係のない会社が担当することも多い。プロジェクトコーディネートの立場から、こういった原則・ガイドといったものを、どのように取り入れていくかは簡単ではない。そうしたケースを前提とした、ユースケースやシナリオを仮想的にでも事例として含まれると非常に参考になる。
  - AI システムは 1 社で完結せずに複数の事業者で成立するという文脈において、関連する事業者は日本国内だけでないことにも留意すべき。
  - スタートアップでは自社でデータを取得することはまれである。ラベリング業務をアウトソースするなど、データ加工はクライアント企業が担当するのが当たり前である。こうなると、企業間の連携は大きな課題になる。
  - 複雑な状況を前提とすると、こうしたらよい書き切ることは難しいと思う。とはいえ、具体的なアクションとして推奨すべきものを記述しておくべき。リスクオーナーをどう定義するかはひとつの論点である。
  - リスク評価は、複数ある選択肢のなかで行われなければいけない。同じような効果が出る手段があるならば、それと比較する。効果が違っても、代替手段としてなにがあるかを考え、そこに対するリスクの差分を見る。実務ではそのようになることが多い。こうした点をガイドラインに記載してもよいのではないか。
  - 当然、発注側はリスクを最小限にしたいと考える。そうなると、結果的にスタートアップを排除するようなガイドラインになってしまうことは危惧される。
  - リスクとリターンのバランスの議論である。リスク管理の側からだけ見れば、負うリスクはゼロが良い。リスクを許容するという考え方とベネフィットをとるという考え方があり、そこはビジネス判断になる。企業ではリスク管理部があり、今までと目線を変えなければ、結果的にスタート

アップを排除するような形になる。本ガイドを使うことで付加価値が上がると理解してもらわなければ利用につながらない。

- **「ゴールベース」の記述の理解に関する論点**
  - 資料中の「ゴールベース」という意味を理解したい。実務の現場では AI は人に対してどういう役割となるべきかという議論がされている。
    - ガバナンスイノベーションは、各社でなにをするかという話よりも、規制レベルという話である。ガイドラインとしてと、社会的なルールとしてという点で、認識がずれてくると考える。「ゴールベース」というのは、AI 原則のように、緩やかに公平性などのゴールが設定されることで、なにがあるとよしであるというブライトラインの考え方ではない。官民が共同にモニタリングしいき、少しずつ良くしていこうという考えである。

- **AI 人材の不足に関する論点**
  - 人材面については課題を感じる。AI について、技術から法律までをカバーできる人材はなかなかいない。透明性といっても、モデルが適当か、という話は非常に難しい。監査の必要性などでは、痛感しているところである。

- **差別等の基本的人権に関する論点**
  - リスクの評価に、差別的なものについて言及してもよい。エンジニアに男性が多いために女性に差別的な事案が生じたなど。また、具体的な例を示すまでいかないが、人の命に関わるとか、基本的人権に関わるとか、そういうときには違う発想が必要だという指摘が必要である。AI 原則のなかでは基本的人権に関する論点が強調されているそうした点について言及があってもいいのではないかと思う。
  - 日本全体のなかで、人権リスクをどう位置づけるかという話がされていないのではないか。もっと上の会議で整理をしなければ先に進めないのではないか。

- **ガイド全体の構成・記述の方針に関する論点**
  - ガイド全体として繰り返しが多く読みにくい印象を受けた。実装・運用段階など、共通しているところがある。ガイドラインの中に列挙されているが、会社の実務担当者としては、まずなにをすべきかが記載されていると読みやすい。どういう段階があるかについては、ケースバイケースなので、触れられていればあとはどの段階で何をするかは担当者が考えればよい。また、チェックリストまでいかずとも、どのようなことに気を付けなければいけないかが挙がっていると参考になる。そのなかにプリンシパルなどがある。例えば、プライバシーに気をつけなければいけないとか、会社に考える材料を与えるという観点でリストアップされているとよい。
  - ガイドラインは網羅的にまとまっているが、読みやすさという点では、マネジメントがやるべきこととプロジェクトがやるべきことが明確に分かれているべきである。
  - AI ガバナンスがある企業ではこうあるべきというゴールが示されているが、マネジメントもプロジェクトも、一気にそのレベルまでいくことはできない。ガイドを読んだ人が尻込みしてしまわ

ないよう、段階的にこう進んでいくということが前段にも書いてあってわかるとよい。

以上