# 経済産業省 第七回 AI事業者ガイドライン検討会
# 議事要旨

令和 7 年 11 月 20 日(木)
16:00～18:00
オンライン会議

## 冒頭挨拶

- 7 月、「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」（以下、AI 法）が制定された。これにより、研究開発及び社会実装を促進する初の包括的な制度枠組みが確立された。
- AI 事業者ガイドラインは、AI 法の基本方針を実務に落とし込む補完的指針として、事業者にとって重要な位置づけにある。一方、知的財産等の関連領域の議論が並行し、内容・分量が過度に肥大化している。
- 生成 AI のマルチモーダル化や各国の政策動向を背景に、当該領域を取り巻く環境変化は加速している。
- 日本は AI ガバナンスの先進国として海外からの照会が増加している。国内では分野別ガイドラインの整備が進む一方、課題に関する意見も多数寄せられている。
- 総務省と連携し、本年度はガイドラインのユーザビリティ向上、多義的な用語の整理、事例の拡充を重点的に検討する。委員各位の闊達で忌憚のないご意見を賜りたい。

## ◆ 全体討議

### ■ ユーザビリティの改善

- **検討に際した留意事項等**
  - ユーザビリティ、活用ガイド、ケーススタディを一体的に検討できると良い。
  - AI ガバナンスに関心を持つ企業を 4 社程度の単位で招集し、課題共有のワークショップを開催している。この取組により、企業の AI ガバナンスに関する問題意識に関する知見が蓄積している。必要に応じてお声掛けいただきたい。
  - 活用ガイド作成過程では、AI 事業者ガイドラインに関するガバナンス・リスクを含む本質的内容の欠落防止を留意していただきたい。
  - 対象者像を明確化した上で作成を進めることが重要であると考える。
  - 活用ガイドが難解で抽象的記述に終始すれば、利用されない事態も懸念される。
  - 他事業者の動向に流され、事業者が AI 関連サービスを不適切に利用する事例が見受けられる。これを踏まえ、当該ガイドの参照により、事業者による不適切なサービス利用の防止を促していただきたい。
- **検討に際した参考文献**
  - 分かりやすさを重視した他の文書では、東京都『文章生成 AI 利活用ガイドライン』および東京商工会議所『中小企業のための「生成 AI」活用入門ガイド』が挙げられ、参考になり得る。
  - 金融庁が本年 3 月に公表した『AI ディスカッションペーパー』が参考になり得ると思料する。金融業界で重要視されるユースケースについて、従来型 AI と生成 AI

の利活用の違い、両 AI サービス間のリスクの差異を体系的に理解しやすい構成となっている。

- **<u>他システムを用いたユーザビリティ改善</u>**
  - チェックリスト相当のツールを用意し、AI にチェックを実行させる方法も考えられる。提供形態として、ファイルおよびプロンプトのみを配布し、システムは各事業者で用意する方式も考えられる。
  - 自社の現状レベルを簡易診断するツールを用い、その結果を踏まえて活用ガイドの使い方を検討していただく、というアプローチも一案であると考える。
- **<u>活用ガイド記載内容</u>**
  - AI ガバナンスを理解する人材の育成と、取組を実行可能な体制の構築が必要である点を明確化することが重要ではないか。
  - ライトユーザーを対象とする場合、ガイドの導入部に「AI ガバナンス構築ステップの全体像」を位置づけることが重要になると考える。活用ガイドの利用促進のため、当該項目には次の三点を盛り込む必要がある。1.ガイドラインを難解と感じるライトユーザーでも導入部で全体像を把握できること。2.自社にとっての必要性を認識し、AI ガバナンス構築を自分事として捉えられるようにすること。3.着手を後押しする内容・構成とすること。
  - ガイドラインのユーザーは、自身に関係するリスクと ToDo を主たる関心事とする。改善は進んでいるものの、活用ガイドで両者の繋がりを整理することが、ユーザーにとって望ましいと思料する。
  - AI 関連ガイドラインが乱立しており、企業は ISO/IEC 42001 や ISO/IEC 5259 等の国際規格に注目している現状にある。事業内容や事業展開地域等に応じて準拠すべき規格・指針を示す記載を、活用ガイドに盛り込むべきだと考える。
  - ライトユーザー向けには、生成 AI と生成 AI 以前の AI で区分した説明が分かりやすいと考える。現行のガイドライン本編では、偽・誤情報、ハルシネーション、著作権を生成 AI 固有に近いリスクとして記載しているが、AI による自律的な決定や、プロファイリングに関わるリスクは生成 AI のリスクとしては考えにくい。活用ガイドのメイン読者として想定する小規模事業者に向け、上記区分による差異を強調すべきと考える。
  - 中小企業を活用ガイドの主たるユーザーに設定する場合、業務効率化等のユースケースを提示することが親切であると考える。
  - 利用シーンごとに、生成 AI とそれ以外の AI の二つに分けて記載するのが妥当と考える。
  - 機械学習においてリスクを完全に除去することは現実的ではない。公平性等のリスク対策をどの程度まで実施するべきかという論点について、可能な限り各事業者の自助努力によりリスクを低減し、致命的リスクの洗い出しとその除去だけでも実行するというリスクアセスメントの発想を活用ガイドに記載するのが良いと考える。

## ■ 多義的な用語の整理

- **<u>開発に関する定義の見直し</u>**

- ベースモデルの短期更新が常態化する現状を踏まえると、実務でのファインチューニングの実施機会は減少傾向にあるとの印象である。むしろ、単一の AI システム内で複数モデルを稼働させるケースの増加を前提に、用語定義の整理を進めるべきではないか。
- 現行の書きぶりは旧来の機械学習の視点が強く、事業者の関心は生成 AI のエージェンティックな利用方法に向いている。これを踏まえ、当該視点を強化するフレーミングとすることが望ましい。
- 出力傾向の制御は、ファインチューニングではなく、プロンプト入力時または生成 AI の出力時のフィルタリングが一般的と考える。モデルの入出力制御という要素を記載すべきではないか。

➢ **<u>データ種類に関する定義の見直し</u>**

- 今の時代に重要なのは、モデル自体の改善ではなく、モデルを社内データにアクセス可能にする方法である。当該点に重点を置く記載が必要であると考える。
- データ定義の追加に関して、各種情報源からの取得方法等、データセキュリティに関する議論の活発化には同意する。一方、訓練用データの取り扱いに関する基礎知識をガイドラインに盛り込む余地がある。AI の対象領域は多岐にわたり、同一用語でも意味が異なり得ることを踏まえ、「データ」の意味合いをより精緻に言語化する必要がある。
- テストデータの項目については、データは指標ではないため、「指標の基になる」への修正が適切であると考える。さらに、「公平な」の記載は、ガイドラインで公平性という用語が厳格なルールの下で使用されていることを踏まえ、「適切な」程度の表現に留める必要があると考える。
- 機械学習におけるデータの観点で表記見直しを検討しているが、生成 AI におけるデータにも言及が必要と考える。

➢ **<u>主体区分の整理</u>**

- ガイドラインユーザーは、自身が開発者・提供者・利用者のいずれに区分されるのか判断に迷うケースがある。インターフェースの定義およびそれに関わる開発者・提供者の役割分担の進展に伴い、開発者兼提供者に該当するユーザーが増加すると予想する。論点は、定義により各主体とリスク、リスクと ToDo がいかに結び付けられるかであり、当該点について一定の見解を示す必要がある。現時点で主体・リスク・ToDo の関係性は不明瞭であるため、対応すべきリスクの漏れを防ぐべく、各概念の整理を進めるべきだと考える。
- 主体区分の定義は、国際動向を踏まえて検討すべきである。他国では、自社でファインチューニングしたモデルを利用する事業者を開発者かつ利用者と解釈する例がある。このようなケースの主体解釈を説明に加えることは有用だが、区分自体は複雑化させない方が良いと考える。
- AI の開発を掲げる事業者の大半は、モデル開発ではなく、RAG 等の既存モデルを用いたアプリケーション構築に従事しているとの印象がある。これらの事業者はガイドライン上で AI 提供者に該当する一方、一般には AI 開発者と呼称される。こうした事業者の実態を精緻に把握する必要がある。
- AI システムの構築・検証フローは、現行のフロー図で開発者に包含しているが、当該方針の見直しが必要であると考える。提供者側が検証を担わない前提では、外部アプリケーション提供者が有害なアプリケーションを提供した場合でも責任追及が困難となり不合理が生じる。また、モデルのみを開発しその後の提供に

関与しない開発者は、モデルのセキュリティ監査は可能でも、具体的ユースケースの監査は困難である。モデル監査とアプリケーション監査を峻別した上で、開発者と提供者の役割の棲み分けを一層精緻化すべきだと考える。

■ **リスクの見直し**

- オープンソースモデルの派生モデルが多数流通している。これら派生モデルの中には、ユーザーの利用中に悪意ある挙動を示すものが登場し得る。モデルの由来の確認方法および利用時の留意点等に言及するのが望ましいと考える。
- 利用者が意図せずフェイクニュース・フェイクデータの生成に関与する事例が実際に発生している。これを防ぐため、モデル開発者・提供者が何に注意すべきか、事案発生時に責任追及が生じ得るか等について、対策やガイドラインを検討していけると良いと考える。
- 各生成 AI プラットフォーマーの利用規約は相互に異なり、実務では見落とされがちだ。同一プラットフォーマーでも利用形態により規約が異なる場合がある。企業データを生成 AI と連携する上で重要な論点であるため、検討対象に含めるのが良いと考える。
- サイバー攻撃により企業活動が妨げられる事案が複数生じている現状を踏まえ、サイバーセキュリティに関する内容を盛り込むのが望ましい。具体的には、一般的なサイバー攻撃に加え、AI/生成 AI による攻撃の増加および特有の問題についても記載すると良いのではないかと考える。
- 他国ではコンパニオン・チャットボット規制の動きが話題となっている。これを踏まえた事例の追加が望ましいと考える。
- AI の活用が進むにつれ、雇用問題などの人権リスク等も増加している。リスクの見直しを実施すべきである。
- 個人が個人へ AI を提供する場合、利用する個人による不適切な利用が増えている。AI 利用者および AI 事業者の責任の在り方についても、検討に含めるのが良いのではないか。
- AI を活用しない、あるいは活用できないことによるリスクへの言及が必要と考える。人手不足が深刻化する中、特に生成 AI を活用しない／活用できないことは、企業の競争力および経営戦略に対する重大なリスクになり得るため、経営層への意識付けを促す必要がある。
- 従前は多様な業務経験を通じて人材を育成してきたが、生成 AI による業務効率化の進展により、社員に付与できる実務経験の機会が制約される可能性がある。人材育成手法の本格的な見直しを怠ること自体が大きなリスクとなり得る点を明記すると良いのではないか。
- リスク対策手法はガイドライン別添で言及されているものの、説明の粒度が粗いと感じる。今後、別添の該当記載をどのように修正していくのか、注視している。

■ **最新技術動向の反映**

- AI エージェントの普及に伴い、AI エージェントと人間の責任分界点の考え方は現状不明瞭である。ガイドラインに、責任分界の考え方に関する指針、想定されるトラブルとその対策の記載を盛り込むのが良いと考える。
- AI エージェントに関する記載を追記する必要がある。
- 消費者向けに AI エージェント的機能の提供が増加しており、日本も例外ではない。商品購入など権利義務に直接影響を及ぼす場面では、契約の有効性等、十分に議論されていない論点が顕在化すると見立てる。
- 消費者向け AI エージェントを提供する場合、AI が消費者本人の情報を自動的に第三者へ提供・入力する場面が増えると見立てる。本人情報の管理が今後の問題になり得ると考える。
- AI エージェント作成のための AI サービスには、独自の検討事項・論点が存在すると考える。近年、ノーコードで AI を作成可能なサービスが拡充しており、ノーコードで作成された AI を作成者不在の状態でも適切に取り扱うために必要な社内ノウハウをいかに蓄積するかという問題が生じると考える。併せて、システムをノーコードの状態でいかに修正していくかという問題も生じ得ると考える。
- 先端 AI 技術の利活用に伴う具体的リスクは不明瞭であり、目まぐるしい技術進歩により予期せぬリスクの発生も想定される。現時点で網羅性の確保は困難との認識の下、先端 AI 技術に関するリスク整理を進める必要があると考える。

## ■ その他

- 消費者間でも AI の利用が一般化し、消費者レベルでの利用方法が話題となっている。今後、消費者が事業者を選択する際の判断材料として、事業者の AI 利活用の形態が加わると予想する。事業者との取引やサービス利用に際し、事業者の AI 利活用の取組を理解することは消費者の責務と位置づける。インターネット通販の需要増を踏まえ、事業者の所在・連絡先・個人情報の取得形態・セキュリティ態勢の確認と併せて、AI 利活用方針を消費者が確認できる制度設計が必要である。事業者がガイドラインを遵守する前提で、消費者にも理解可能な内容・構成へ修正すれば、事業者選択の際の参考となると考える。
- 現行ガイドラインは、著作権および個人情報保護法に関する記述が少ない。特に中小企業向けの追記を検討する余地があると考える。

以上