

経済産業省委託事業

令和7年度技術開発調査等推進事業

(AI利活用における民事責任の在り方に関する調査研究及び研究会運営)

## 調査報告書

令和8年3月

みずほリサーチ&テクノロジーズ株式会社

「令和7年度技術開発調査等推進事業（AI利活用における民事責任の在り方に関する調査研究及び研究会運営）」は、経済産業省の委託事業として、みずほリサーチ&テクノロジー株式会社が実施したものです。本報告書の内容を引用・転載する際は、出典を明記していただきますようお願い申し上げます。

# 目 次

第 1 章 調査概要 .....	1
1. 実施目的 .....	1
2. 実施内容 .....	2
3. 実施体制 .....	2
第 2 章 研究会実施記録 .....	3
1. 開催概要 .....	3
1.1 構成員 .....	3
1.2 開催状況 .....	4
2. 討議記録 .....	5
2.1 第 1 回 .....	5
2.2 第 2 回 .....	12
2.3 第 3 回 .....	19
2.4 第 4 回 .....	26

# 調査概要

---

## 1. 実施目的

生成 AI の登場以降、事業領域における AI サービスの利活用は年々拡大しており、近年では、AI を搭載した IoT 製品やロボティクス等、高度な自律システムの開発及び社会実装も進みつつある。また、システムの高度化に伴い、開発や運用のための大規模かつ高品質なデータへのニーズが高まり、事業者間の垣根を超えたデータ連携による新たな付加価値の創出や社会課題の解決も期待されている。

このように、AI やデータの高度利用によってサイバー空間のみならずフィジカル空間への作用が増大しつつある中、政府は、AI がもたらす社会的リスクの低減を図るとともに AI のイノベーション及び活用を促すべく、非拘束的なソフトローによって目的達成に導くゴールベースの考え方で「AI 事業者ガイドライン」を取りまとめた。他方、リスクが顕在化し事故が発生した際の民事責任の在り方については、AI の自律性やブラックボックス性を踏まえてどのように解釈適用を行うべきか、裁判例や統一的な見解がなく、責任の所在が不明瞭であることがシステムの導入や開発を躊躇させる一因となっている。また、事業者において「AI 事業者ガイドライン」を遵守していた場合に、そのことが責任論の枠組みにおいてどのように評価されるのかが不明瞭であることも指摘されている。

上記の背景を踏まえて、AI を用いたシステムが事故に寄与した仮想事例を題材に、不法行為法及び製造物責任法の観点から解釈適用上の論点及び考え方の整理を行うことを目的として、本事業を実施した。本事業では、事故発生時の責任論について予測可能性を高めるとともに、「AI 事業者ガイドライン」が責任判断に当たって有益な基準となる場合があることを示しつつ、被害者や企業担当者、裁判官等の関係者に対し論点の所在及び考え方の指針として、「AI 利活用における民事責任の解釈適用に関する手引き（案）」を取りまとめた。

## 2. 実施内容

### (1) 研究会の運営

有識者を構成員とする「AI利活用における民事責任の在り方に関する研究会」を計4回開催し、事務局としての運営に必要な業務（資料等の作成、研究会の事前調整事務、議事録作成等）を実施した。同研究会では、AIが関係する事故等の発生時における民事責任論の解釈適用上の論点及び考え方についての検討を行った。

### (2) 研究会報告書執筆の補助業務

上記研究会の報告書案として「AI利活用における民事責任の解釈適用に関する手引き（案）」を取りまとめるとともに、取りまとめ過程における校閲や図表作成等の執筆補助業務等を実施した。

また、令和8年2月18日から開始された意見公募手続（パブリックコメント）を踏まえて、上記報告書全体の校閲を行うとともに、報告書の概要説明資料を作成した。

### (3) 調査結果の英訳

上記報告書案及び概要資料の英訳を実施した。

### (4) 連絡会議の実施

経済産業省との間で、本業務の遂行において必要な打合せを実施した。

### (5) 調査報告書の作成

上記の業務内容を取りまとめた調査報告書（本書）を作成した。

## 3. 実施体制

本調査研究の実施体制は、下図のとおりであった。

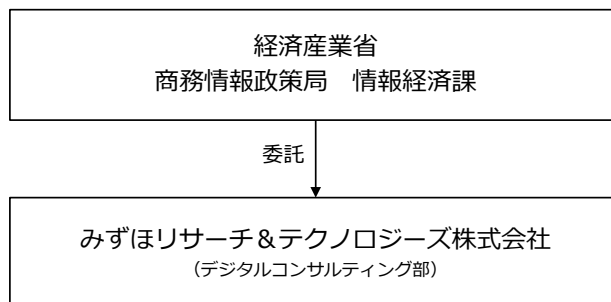


図 1 本調査研究の実施体制

# 研究会実施記録

---

## 4. 開催概要

本業務では、有識者を構成員とする「AI利活用における民事責任の在り方に関する研究会」を計4回開催し、事務局運営業務を実施した。本研究会では、AIが関係する事故等の発生時における民事責任論の解釈適用上の論点及び考え方についての検討を行った。

### 4.1 構成員

本研究会の構成員は、以下のとおりであった。

#### <座長>

大塚 直                      早稲田大学法学学術院 教授

#### <委員> 50音順

稲谷 龍彦                      京都大学大学院法学研究科 教授  
江間 有沙                      東京大学国際高等研究所東京カレッジ 准教授  
柿沼 太一                      STORIA 法律事務所 代表パートナー弁護士  
宍戸 常寿                      東京大学大学院法学政治学研究科 教授  
白石 友行                      上智大学法学部 教授  
中原 太郎                      東京大学大学院法学政治学研究科 教授  
橋本 佳幸                      京都大学大学院法学研究科 教授  
福岡 真之介                      西村あさひ法律事務所・外国法共同事業 パートナー弁護士  
松尾 豊                      東京大学大学院工学系研究科 教授

#### <オブザーバー>

内閣府科学技術・イノベーション推進事務局  
法務省民事局  
法務省大臣官房司法法制部  
総務省国際戦略局  
消費者庁消費者安全課  
AIセーフティ・インスティテュート (AISI)  
一般社団法人日本損害保険協会  
一般社団法人日本ディープラーニング協会  
日本弁護士連合会

## <事務局>

経済産業省 商務情報政策局 情報経済課

みずほリサーチ&テクノロジーズ株式会社 デジタルコンサルティング部

## 4.2 開催状況

本研究会は、以下のとおり、計4回開催された。

表 1 「AI 利活用における民事責任の在り方に関する研究会」：開催状況

開催回	日時 (すべてオンライン開催)	議題
第1回	令和7年8月19日(火) 10:00~12:00	<ul style="list-style-type: none"><li>● 想定事例1について</li><li>● 想定事例2について</li><li>● 複数当事者間の責任判断について</li></ul>
第2回	令和7年10月1日(水) 10:00~12:00	<ul style="list-style-type: none"><li>● 第1回研究会の議論の内容について</li><li>● 想定事例3について</li><li>● 想定事例4について</li><li>● 想定事例4に関する立証上の論点について</li></ul>
第3回	令和7年12月3日(水) 15:00~17:00	<ul style="list-style-type: none"><li>● 第2回研究会の議論の内容について</li><li>● 想定事例5について</li><li>● 想定事例6について</li></ul>
第4回	令和8年1月14日(水) 15:00~17:00	<ul style="list-style-type: none"><li>● 第3回研究会の議論の内容について</li><li>● 報告書案に関する検討</li></ul>

## 5. 討議記録

各回における討議の概要は、以下のとおりであった。

### 5.1 第1回

#### ① 開催日時

令和7年8月19日（火）10:00～12:00（オンライン会議）

#### ② 議事次第

1. 開会
2. 議事
  - (1) 想定事例1について
  - (2) 想定事例2について
  - (3) 複数当事者間の責任判断について
3. 閉会

#### ③ 配布資料

【資料1】議事次第・配布資料一覧

【資料2】AI利活用における民事責任の在り方に関する研究会の開催について

【資料3】事務局説明資料

#### ④ 出席者

大塚座長、稲谷構成員、江間構成員、柿沼構成員、宍戸構成員、白石構成員、中原構成員、橋本構成員、福岡構成員、松尾構成員、オブザーバー（関係省庁・関係団体）、事務局

#### ⑤ 議事概要

##### 1) 本研究会の方針について

資料3の2～6頁に基づき、事務局から本研究会の趣旨・目的及び検討対象について説明が行われた後、構成員による討議が行われた。

##### 【成果物の位置づけや名称について】

- 本研究会の成果物の名称について、資料で記載されている「準則」という言葉は、法学では、適用されるべき規範や基準等の意味合いで用いられることが多い。本研究会の目的が、民法第709条における過失や因果関係等の既存の概念を変えることなく、AIに関する事案への適用結果を検討することにあることを踏まえると、「解釈の指針」や「考え方」等の表現のほうが、より適切ではないか。

- 法律の解釈に関する最終的な判断は裁判所が示すものであるため、本研究会の成果物の位置づけについても慎重な検討が求められる。
  - 本研究会と目指す方向性が類似する文書の例として、資料中で「電子商取引及び情報財取引等に関する準則」を紹介した。しかし、成果物の名称を「準則」とすることに必ずしも固執するものではないため、「考え方」や「指針」等の位置づけの文書にすることも視野に入れ、最終的な成果物の内容を踏まえて、改めて名称を検討したい。

### 【本研究会の検討対象とする事例について】

- 資料中に①（人の行為や判断に補助的に AI を活用する場合）と②（AI が最終的な判断まで下す場合）の 2 つの事例が示されているが、①の類型は、さらに 2 つに細分化できると考えている。1 つは AI が専門的な業務の下準備等の補助的な役割を担い、人間による確認が当然必要と考えられる形態であり、もう 1 つは画像認識 AI のように、ある領域で人間と同等以上の精度を持つと考えられるものである。特に後者の高性能な AI を利用する場合、常に人間による厳密な再確認を義務付けることは、このような AI の普及を阻害する可能性があるため、AI の判断を尊重しながら進めるほうが望ましいと言えるような類型も想定する必要がある。なお、特定の領域で人間を上回る性能を発揮する AI は、現在の技術水準でもすでに現れ始めているため、AI が汎用的に人間を超える能力を獲得する、いわゆる「汎用人工知能（AGI）」の実現や「技術的特異点（シンギュラリティ）」の議論とは区別する必要がある。
- AI が専門的なプロフェッショナルと同等以上の判断を行うケースは非常に重要である。また、AI の利用形態として、意思決定の補助だけでなく、人材育成やトレーニングの場面で AI が教育者の役割を担うケースも増加している。将来的には、このような形で育成された人材の能力担保や品質保証等の問題も重要になると考えられる。

## 2) 想定事例 1 について

資料 3 の 7～10 頁に基づき、事務局から、想定事例 1（通常業務における判断補助 AI の事例）について説明が行われた後、構成員による討議が行われた。

### 【資料 10 頁：類型区分について】

- 資料 10 頁の類型 1（説明義務違反）と類型 2（侵害行為を容易にしたことによる責任）を区分する視点は何か。
  - 現状では、類型 1 を情報提供上の義務違反、類型 2 を現に存在するリスクに対して適切な措置を講じずに放置した責任、と整理している。
  - 類型 1 と類型 2 を区分する視点として、ドライバー B' が AI の判断を鵜呑みに

する危険を、AI 開発・提供事業者 A が積極的に作り出したか、それとも運送事業者 B の下で生じている危険を、AI 開発・提供事業者 A が不作為により放置したか、という違いで整理することも考えられる。本質的に異なるように見えない印象もあり、類型を区別している視点や基準は明示していただけるとよい。

- 配送ルート最適化 AI は、通常、道路の安全性を保証するものではなく、あくまで効率的なルートを示す補助ツールである。したがって、実際の道路状況に応じた安全確保の判断は、AI 開発・提供事業者 A ではなく利用者（運送事業者 B）に委ねられていると考えられる。こうした役割分担は、第三者（配送先 V）との関係でも考慮されるべきであり、結果として、ドライバー B' は AI の出力とは独立に安全な走行経路を選択する義務を負うと整理できる。なお、この義務・責任範囲の振り分けの観点から、資料 10 頁の AI 開発・提供事業者 A の責任を整理すると、類型 1 は安全性が担保されているかのように利用者に誤信させることによって適切な責任の振り分けが行われなかった場合であり、類型 2 はそもそも利用者が安全性を確保できないような、事故を実質的に誘発するシステムを提供した場合と整理できる。

#### 【資料 10 頁：参考判例の位置づけについて】

- 資料 10 頁で AI 開発・提供事業者 A の責任の類型として挙げられている参考判例は、本想定事例と事案の性質がやや異なる。特に類型 2 を「侵害行為を容易にした」とまとめることは適切ではない可能性がある。「侵害行為を誘発した」あるいは「引き起こした」といった表現のほうが、事案の実態に近いのではないかと。
  - 参考判例については、事案が完全に一致するものではないが、ツール提供者の責任がどのように導かれるかを考える上での参考として挙げたものである。「侵害行為を誘発した」あるいは「引き起こした」といった表現は合理的であり、適切な表現を改めて検討したい。
- 本想定事例の解決の方向性に異論はないが、参考とされたカーナビの判例は、利用者が製造者に対して損害賠償を求めた事案であり、第三者である配送先 V との関係が問題となる本想定事例とは構成が異なる点に留意が必要である。

#### 【想定事例 1 の構成及び論点について】

- AI の限界等について適切に警告を行っていたのであれば、原則として AI 開発・提供事業者 A は責任を負うべきではないという考え方が自然であり、かつ、AI の社会実装を不当に阻害しないためにも重要である。一方で、適切な指示・警告がなかった場合には、AI 開発・提供事業者 A の責任が問題となり得る。この場合、AI 開発・提供事業者 A と運送事業者 B の間には、通常、契約関係が存在し、免責規定が設けられていることが多いと考えられる。AI 開発・提供事業者 A に故意・重過失が認められ

ば、免責規定があっても AI 開発・提供事業者 A が責任を負う可能性はあるが、その場合でも明らかに通行困難な道路に進入したドライバー B' の過失を考慮し、過失相殺がなされると考えられる。このように考えると、本事例を通じて示される行為規範は、AI 開発・提供事業者 A は適切な指示・警告を行うべきであるとともに、契約に免責規定を設けることが重要であり、さらには、故意・重過失があれば責任を負う、ということになる。

- 適切な指示・警告を行ったかという点はきわめて重要である。表示の仕方によっては、警告がなされていても利用者にそのリスクが適切に伝わらず、AI に対する過度な信頼を生じさせてしまう可能性がある。このような状況を認識しながら放置し、結果として問題が起きたという類型も考えられるのではないか。
- 本想定事例のような場合は、保険適用に関する論点も考えられる。

### 3) 想定事例 2 について

資料 3 の 11~14 頁に基づき、事務局から、想定事例 2（専門業務における判断補助 AI の事例）について説明が行われた後、構成員による討議が行われた。

#### 【異なる類型設定の必要性について】

- 前述のように、高精度な AI については、そのような AI を信頼して利用することをある程度許容し、利用者の注意義務がある程度軽減されるような類型も検討しなければ、AI の利用が進まないのではないか。また、AI 開発者にとっても、開発した AI が一定レベルの性能を達成していれば、責任を問われないようにすることや、責任の程度を判断する上で、開発体制の適切性なども考慮されるような仕組みがなければ、AI 開発事業者が委縮してしまう可能性がある。このような類型も、今回の事例とは別に議論する必要があるのではないか。
  - 弁護士等の専門職については弁護士法のような規制法があるため、最終的な判断はやはり専門家が行わなければならないのではないかという議論は、民法学者には強いようである。
  - 専門家の責任が AI を信頼することで一切なくなるということではないが、明らかな誤り以外は、AI を一応信頼しても良いという形などで注意義務が軽減されることはあってもよいと考えている。

#### 【想定事例 2 の位置づけについて】

- 本事例は、専門家の注意義務の問題として整理できる。弁護士は依頼者に対して高度な注意義務を負っているため、想定事例 1 のドライバー B' と比較して、AI の出力を鵜呑みにせず、自分自身で確かめる責任はより重いと解される。LLM 提供事業者 A 及

び AI サービス開発事業者 B の責任は、提供相手が高度な注意義務を負っている専門家であることから、想定事例 1 の場合と比較して相対的に軽くなると考えられる。

- 本事例の論点となっている専門業務は、資格者に限ったものではないと考えている。専門業務の要素として挙げられている資料 14 頁の①から③は、妥当なものであると思われるが、それに加えて、さらに依頼者が専門家としての知識を期待してその人に依頼しているという点も基準になると考えられる。つまり、専門性を期待した契約がなされている場合は、専門家としての高度注意義務を負うことになると言える。
- 資料 14 頁の①から③は、業務の性質の問題というより、AI 利用者の専門家性の問題として位置づけたほうがよい。
- 想定事例 1 と 2 の違いは、「専門業務」であるかどうかという点であるが、資料 10 頁の表では、専門性以外の要素も挙げられており、専門性だけを強調した枠組みが適切な点かという点が懸念される。
- 想定事例 1 と 2 は、いずれも利用者が最終判断を下すことが前提である点で、構造上の違いはそれほど大きくない。両者の違いは、責任判断の枠組みそのものではなく、求められる注意義務の程度の差と捉えるべきではないか。

#### 【資料中の表現等について】

- 資料 14 頁に、「幫助責任が成立する余地もある」という記載があるが、幫助責任は限定的に解釈されるべきである。そもそも専門家が自分で判断すべきであるのに、AI 開発者も責任を負う余地があるように見える表現は望ましくない。  
→ 当該箇所「限定的に」という表現を加えることも考えられる。
- 想定事例 1 の AI は期待どおりの出力をしているが、想定事例 2 の AI は、期待された結果を生み出していない。それでもなお弁護士 C が責任を負うのは、弁護士 C が職業上の法的義務として、AI 利用の有無にかかわらず、自らの判断で法的サービスを提供する責任を依頼者 V に対して負っているためである。この違いを踏まえれば、資料 14 頁下線部の、責任範囲が「通常業務と比較すれば狭まり得る」という記述は理解できるが、「専門業務において用いられる AI は、開発・提供における配慮が少なくてもよい」という誤解を招かないように、表現には留意が必要である。例えば“説明義務違反あるいは幫助責任として評価される場合が少ない”という趣旨の表現のほうが望ましいのではないか。

#### 4) 複数当事者間の責任判断について

資料 3 の 15～20 頁に基づき、事務局から、AI サービス提供に係る主体が複数にわたる場合における論点について説明が行われた後、構成員による討議が行われた。

### 【事例の構成について】

- 依頼者 V は、直接の契約相手である弁護士 C の責任を迫ることが通常であり、LLM 提供事業者 A から AI サービス開発事業者 B への説明の十分性等の内部的な事情は依頼者 V からは見えにくい。したがって、LLM 提供事業者 A や AI サービス開発事業者 B の責任が問題となるのは、主として、弁護士 C が依頼者 V に賠償した後の、弁護士 C から LLM 提供事業者 A・AI サービス開発事業者 B に対する求償関係においてであろう。その際は、契約で予定された性能が満たされていたか、LLM 提供事業者 A・AI サービス開発事業者 B からの説明が十分であったか等の点が争点になると考えられる。  
→ 因果関係は認められる場合はあるので、直接 LLM 提供事業者 A・AI サービス開発事業者 B に対して請求する途についても検討していきたいが、今のご指摘については、どこかで言及したい。

### 【説明義務について】

- AI の出力は、学習用データ、ユーザーが入力したプロンプト、RAG 等で参照される外部データベースの3つの要素によって規定される。したがって、説明義務の範囲も、これらの要素を基に考えるべきである。AI サービス提供者は、どのようなデータで学習し、どのデータベースを参照することで、どのような範囲の判断が可能で、どのような限界があるのかという事項を説明する義務を負っていると言える。
- 説明義務について、例えばまったく学習していないことを学習したと説明するような虚偽の説明は、説明義務違反になるということは言えるが、何を学習させたかといったような具体的な説明の範囲や水準を一律に定めることはきわめて困難であり、最終的にはケース・バイ・ケースの判断とならざるを得ないのではないか。一つの考え方として、その業界における一般的な知識に照らして、説明が妥当であったか否かを判断するというアプローチが考えられる。

## 5) 全体を通して

### 【座長総括】

- 本日の意見を踏まえ、想定事例 1 については、指示・警告がある場合とない場合に分け、また、ない場合については、免責規定がある場合とない場合に分けて検討することが重要である。
- 想定事例 2 について、専門業務と通常業務に分けるよりも、各事例をフラットに扱ったほうがよいという意見もいただいたが、分かりやすさのためにこのように分けることは重要であると考えている。
- 信託の原則が適用される交通事故とは問題状況が異なっており、専門家については、

AI を信頼したとしても注意義務が軽減されにくいと言えるが、類型をさらに分けて、注意義務が軽減される可能性を検討したほうがよいという指摘もあったため、引き続き、検討を行いたい。

- 複数当事者の場合の説明義務についての議論の中で、業界の標準についての指摘があり、非常に重要な点であると感じた。また、実際には、複数当事者が関わる事案において、依頼者 V は弁護士 C に責任を追及することが多いが、AI サービスの提供側である LLM 提供事業者 A や AI サービス開発事業者 B に対して責任を追及する可能性もあり得るため、検討はしておいたほうがよいと考えている。

## 5.2 第 2 回

### ① 開催日時

令和 7 年 10 月 1 日（水）10:00～12:00（オンライン会議）

### ② 議事次第

1. 開会
2. 議事
  - (1) 第 1 回研究会の議論の内容について
  - (2) 想定事例 3 について
  - (3) 想定事例 4 について
  - (4) 想定事例 4 に関する立証上の論点について
3. 閉会

### ③ 配布資料

- 【資料 1】 議事次第・配布資料一覧
- 【資料 2】 構成員等名簿
- 【資料 3】 事務局説明資料

### ④ 出席者

大塚座長、稲谷構成員、柿沼構成員、宍戸構成員、白石構成員、中原構成員、橋本構成員、福岡構成員、松尾構成員、オブザーバー（関係省庁・関係団体）、事務局

### ⑤ 議事概要

#### 1) 第 1 回研究会の議論の内容について

資料 3 の 3～4 頁に基づき、事務局から前回研究会の討議概要と対応方針について説明が行われた。

#### 2) 想定事例 3 について

資料 3 の 5～14 頁に基づき、事務局から想定事例 3（画像認識 AI の事例）について説明が行われた後、構成員による討議が行われた。

#### 【AI 利用者の過失責任に対する考え方について】

- 人間が高度な AI と協調的に作業を行うと、人間の注意水準は低下するという認知科学的な知見を踏まえると、一定水準以上の精度を持つ AI の利用時に、利用者が AI の検出漏れを発見できなかったことをもって、直ちに利用者の過失を問うことは望まし

くない。海外では、AIによる適切な診断結果を医師が受け入れなかった結果、治療が遅れて患者に不利益が生じ、訴訟につながった例もあるとのことである。AIを信頼して利用しないとかえって危険な可能性も生じ始めており、そのような方向での議論が望まれる。

#### 【AI利用時の注意義務の質的転換について】

- 資料の検品受託事業者Bの注意義務について、基本的に異存はない。ある作業においてAIを利用する場合、注意義務の対象は、参考裁判例③にも見られるように、人間の作業からAIの適切な作動を確保することへと変化する。AIを利用することで、人間が作業する場合と比べて注意義務が軽減されるわけではないと考えられるため、取りまとめの際は注意していただきたい。
- 人が判断を行う場合は分岐a及びbのように個別の事例における異物の発見可能性が問題になることに対し、AIを利用する場合はシステムの適切性が問題となるのであり、過失の次元が異なる点を強調したほうがよい。
- 分岐aと分岐bという区分は、第1回研究会の想定事例1や想定事例2には当てはまるが、想定事例3を分析する枠組みとしては、必ずしも合致しないように思われる。
- AIを利用することにより、人間による作業と比べて、過失の判断の質的な転換があるという点と、質的な転換があるとはいっても、これまでの過失の議論の延長線上で捉えられるという点は、明確にしておいたほうがよい。
- AI利用時における注意義務の質的転換は無条件で認められるわけではなく、AIの精度が十分高いことや全件を手で確認することが困難等の事情により、AIの利用に合理性が認められることが必要になると考えられる。
- 注意義務に関しては、AI利用時の体制構築義務の中に、使用が認められるAIの性能水準という観点が含まれるかという点や、当該領域で精度の高いAIが一般的に利用される状況であった場合に、そのようなAIを使用することにまで義務が及ぶのかどうかといった点も論点になり得る。

#### 【想定事例3の内容や位置付けについて】

- 「AI利用者」の定義が不明確な印象を受ける。想定事例3における製造業者Cも、AIを用いた検品サービスを利用するという意味では「AI利用者」であるが、製造物責任という重い責任を負うため、資料14頁の表を報告書に記載する際は注意が必要である。
- 想定事例3において、検品受託事業者Bは、製造業者Cから引き受けた範囲でのみ検品の責任を負うため、検品受託事業者Bの行為義務の水準はあまり高くないと考えられる。このため、AIの過誤の出力について、検品受託事業者BやAIサービス開

発事業者 A の無過失責任を検討する等の議論の一般化に用いる場合は、注意が必要である。

- 製造業者 C が検品受託事業者 B に委託せず、自社で AI を用いた検品を行う場合も当然想定され、このような場合、製造上の欠陥として過失判断を介さずに製造業者の製造物責任が生じ得る。本事例に関する議論が、既存の過失判断の枠組みの延長で捉えられることを明確にする意味でも、この点について注記すべきである。
- AI サービス開発事業者 A が複数の類似分野の事業者 서비스에提供している場合など、必ずしもユーザーである検品受託事業者 B より検品対象等に関する知識に乏しいとは限らないというケースも想定する必要があるのではないか。

#### 【資料 14 頁：想定事例 1 から 3 の整理について】

- 資料 14 頁の想定事例 1 から 3 の整理について、従来の枠組みに沿った整理が可能である。一般的に、過失は注意義務と結果回避義務を果たしていたかという点で判断されるが、通常、この結果回避義務と注意義務は一般人基準で判断される。ただし、専門家の場合は、専門家としての注意義務と専門家としての結果回避義務があると考えられると、想定事例 1 から 3 も横断的に捉えることが可能であり、従来の過失の議論とも整合する。分岐 b についても、精度が高いというだけでなく、精度が高いと、注意義務のほか、損害発生の予見可能性や結果回避義務が低減され得ると捉え、それらが過失判断において考慮され得ると考えれば、従来の過失論の枠組みと整合的に整理できるのではないか。
- 資料 13 頁の AI サービス開発事業者 A の注意義務について、異存はない。資料 14 頁の分類について、A) に記載されている AI 開発者の義務は、B) でも妥当する。また、B) に記載されている AI 開発者の義務の範囲が、AI 利用者との合意によって規定されるという内容は A) でも妥当するため、この A) と B) の対比は、ややミスリーディングであると思われる。
- 資料 14 頁右側の 4 つの考慮要素は、一律に当てはめるのではなく、場面に応じて使い分けることが重要である。また、欧州では、プロファイリングに使用される AI については、AI の自動化された判断に服さないことが想定されていることを踏まえると、②の「リスクの性質・程度」に関わる要件が「生命、身体」のみで十分か、検討が必要である。さらに、④の「便益」という用語に含められる意味が多様であるため、あまり広がらないように捉えていくほうがよい。

#### 【専門家の注意義務について】

- 資料 14 頁右側の判断基準は、想定事例 1 と 3 では当てはまるが、想定事例 2 では当てはまらないように思われる。仮に AI の精度が向上して AI の判断を尊重する義務の

ようなものが生じ得るとしても、医師や弁護士のような専門家においては、最終的な判断責任を自身が負うという基本線は揺るがないと考えられるため、一般の注意義務と同列に論じることはできない。

- 認知科学の知見からは、トレーニングを積んでいる人であっても AI が高度な正確性を発揮した場合の注意水準の低下は避けられない。当該領域において AI をどう活用すべきかを判断する「適切な利用方法に関する高度な能力」を、専門家としての注意義務の中心とすることが合理的ではないか。

#### 【確率的に動作する AI と因果関係について】

- 想定事例 3 において、仮に AI の精度が不足していたというシステム構築上の過失が認められたとしても、その結果は確率的に出力されるものであって、当該過失と特定の人物に対する個別の被害発生との間の因果関係を証明することは困難である。
- AI により確率的に発生する事象に対応するため、新たな因果関係の捉え方を検討することが今後の重要な課題となる。刑法分野では、AI 開発者又は利用者にリスクを一定水準以下に低減する義務を負わせ、当該義務への違反が生じている状態で事故が発生した場合には、それを注意義務違反により増加した危険が具現化したものと解釈するというような議論も行われている。
- 今後、確率的因果関係のような考え方にに基づき、割合的に責任を認めていくという選択肢もあり得るのではないか。

### 3) 想定事例 4 について

資料 3 の 15～23 頁に基づき、事務局から、想定事例 4（画像生成 AI の事例）について説明が行われた後、構成員による討議が行われた。

#### 【事例の構成や設定について】

- 想定事例 4 の整理については、以下に指摘されている点を除き、おおむね異存はない。
- 開発者の注意義務自体は、想定事例 1～3 と本質的に異なるものではない。予想される利用方法に照らして他人の法益侵害の可能性を合理的に低減する開発及び必要な付随措置を講じるという通常の過失判断の延長として整理が可能である。
- 最高裁平成 24 年判決の判示との関係では、権利侵害（パブリシティ権侵害）及び違法性に加えて、故意・過失の有無を問うことが必要であるという説明が必要であるように思う。
- ツール提供者の主体责任と幫助責任に分けた議論が行われているところ、著作権法を参考にしていただいていると思われるが、差止請求については明文規定のある著作権法と異なり、

これが認められるか、また共同不法行為の際にどのように考えるかなどの論点は、一意に決められるものではない。

- 著作権を念頭に検討されている印象を受けるが、パブリシティ権は、複製の時点では侵害とならない点や私的使用として認められる範囲が広い点など、著作権とは性質が異なる。パブリシティ権や肖像権は、著作権に比べて権利侵害となる利用態様が狭いため、不法行為が成立すると判断される場合も少なくなるという評価も可能と考えられる。
- 資料 22 頁の分岐 b について、AI 開発ベンダ D 社による AI 提供行為が単独でパブリシティ権を侵害するという整理によると、AI 開発ベンダ D 社の不法行為を、アパレル業 E 社が画像を商業利用することによる侵害行為とは別に考えることになる。他方、参考裁判例⑧及び⑨では、1 つの著作権侵害行為があり、その主体が誰かということが問われているため、完全に対になるような形で整理してよいか気になった。  
→ 分岐 b におけるアパレル業 E 社の行為が資料中に明示されていないため、画像出力自体を問題とするのか、その後の商業利用を問題とするのかが明らかでない。アパレル業 E 社の行為は明示したほうがよい。
- 資料 20 頁について、結論として異存はないが、参考裁判例⑩のカラオケ装置は、利用者による権利侵害行為を生じさせる蓋然性がかなり高いケースであるのに対して、画像生成 AI は使い方に依存する側面があるため、これと同様に整理をしてよいかは検討が必要に思われる。
- 権利侵害防止措置を取るべき目安として、データセットの規模が挙げられている点に違和感はないが、具体的な規模については、技術的妥当性の確認や検証が必要である。
- データを学習させた時点では無名だった人物が後に著名人になった場合など、時間経過の中で権利侵害リスクが発生した場合に、どの主体がどのような注意義務を負うかという点についても整理が必要ではないか。継続的に提供されるサービスであれば、契約で適切なアップデートを行うことを定めることも考えられるが、買い切り型の製品では深刻な問題となる可能性がある。

#### 4) 想定事例 4 の立証上の論点について

##### 【過失の判断について】

- 過失の択一的・概括的認定と包括的推定は区別が必要であり、過失を構成する義務自体の特定が択一的にも概括的にも困難な場合にまで推定を広げることについては、慎重な検討が求められる。
- 今回は想定事例 4 についての整理であるが、想定事例 3 における検品受託事業者 B の過失判断においても、検品受託事業者 B が用いた AI の精度や講じたリスク低減措置の内容は被害者（消費者 V）からは把握が難しいため、立証上の配慮に関する検討が

必要となり得る。

#### 【実効的な証拠収集について】

- 文書提出命令の相手方が訴訟当事者である場合と第三者である場合で、要件や運用に差異が生じるかという点については論点となり得る。
- 営業秘密を理由とする不開示が想定されるため、特許訴訟におけるインカメラ手続のような仕組みの導入を立法論として検討することが考えられる。実務上の守秘義務契約等の補完策もあるが、制度的裏付けがないと実効性に欠ける。
- 証拠を出すよりは訴訟に負けて、あるいは和解によって金銭で解決するといった方が得だから証拠を出さない、ということが多くなると、結局、安全性を高めるための情報は集まらなくなってしまう。行政による規制・監督のための証拠に関する情報収集の仕組み等、民事訴訟以外の制度との連携も視野に入れて、社会全体として AI の安全性を確保する仕組みの設計が重要である。

#### 5) 全体を通して

##### 【座長総括】

- 想定事例 3 では、過失判断の焦点が「個々の検査結果の確認」から「AI が望ましい作動をするためのシステム設計・運用」に移るとの理解が広く支持された。これは、過失の内容が変わったということであり、「注意義務が軽減した」ということではないため、そのような表現や位置づけは適切ではないという指摘があり、重要である。
- 資料 14 頁について、従来の予見可能性や結果回避義務の枠組みで、高精度 AI が過失判断に与える影響についての整理が必要であるとの指摘があった。
- (検査について) システムを用いること自体の適否の問題があり(注意義務の質的転換)、検査の精度が高いか否か等を基準とする資料 14 頁の分類は、システムを用いること自体の適否を判断する基準と捉えることができるとの指摘は重要である。そして、これを、AI を用いることの適否についての過失の判断基準として用い得るという指摘は、納得感のある考え方である。
- 想定事例 4 では、時間軸に関する指摘を踏まえ、AI 開発ベンダ D 社がどの時点でのどのような注意義務を負うかという点について整理が必要であるとの指摘があった。過失はそれぞれの時点で判断すべきであるが、開発者にアップデートの義務がある場合は D 社の過失が問題となり得ると考えられる。
- AI により、専門家の注意義務の内容がシステムの使い方・運用に変わっていくのではないかと指摘があったが、民法においては、最終的な判断の責任は専門家が負うべきという考え方がやはり主流と思われる。

- 「AI利用者」の定義が不明確との指摘があった。製造業者 C 自身が AI を用いる場合も「AI利用者」となり得るため、製造物責任の扱いを含めて、検討が必要である。
- 想定事例 4 では、故意・過失の記載を補強する必要があるとの指摘や、D 社に裁判例におけるカラオケ装置のリースの事業者ほどの関与があるかは明らかでないとの指摘があり、報告書での表現に反映したい。
- 立証論について、文書提出命令に関して訴訟当事者か第三者かの区別が必要であるという点や、インカメラ手続の必要性を言及すべきとの点は重要である。
- AI の適正化・安全性確保に向けた情報収集のために、他制度との連携により、情報が社会に還元される仕組みを実現することが必要であるとの指摘があった。極めて重要であると思われる。
- 因果関係に関して、個別的因果関係の立証が困難な AI 特有の問題について、AI の確率的な挙動と因果関係との関係をどのように考えていくか、問題提起があった。

## 5.3 第3回

### ① 開催日時

令和7年12月3日（水）15:00～17:00（オンライン会議）

### ② 議事次第

1. 開会
2. 議事
  - (1) 第2回研究会の議論の内容について
  - (2) 想定事例5について
  - (3) 想定事例6について
3. 閉会

### ③ 配布資料

- 【資料1】 議事次第・配布資料一覧
- 【資料2】 構成員等名簿
- 【資料3】 事務局説明資料

### ④ 出席者

大塚座長、稲谷構成員、江間構成員、柿沼構成員、宍戸構成員、白石構成員、中原構成員、橋本構成員、福岡構成員、松尾構成員、オブザーバー（関係省庁・関係団体）、事務局

### ⑤ 議事概要

#### 1) 第2回研究会の議論の内容について

資料3の3～4頁に基づき、事務局から前回研究会の討議概要と対応方針について説明が行われた。

#### 2) 想定事例5について

資料3の5～19頁に基づき、事務局から想定事例5（取引審査AIの事例）について説明が行われた後、構成員による討議が行われた。

#### 【本事例の位置づけや枠組みについて】

- 結論に異論はないものの、全体の構造がやや分かりにくい。従来は人間が判断主体であることを前提に、取引拒絶の判断の決定的な理由を特定した上でその合理性を判定してきた。これに対して、AIの判断過程はブラックボックスであり、取引を拒絶する理由の特定が困難であるため、統計的なバイアスや公平性の指標の乖離を違法性判断

の根拠とせざるを得ない。判断を行った AI の動作ではなくて、あくまで開発者によるコントロールや利用者によるチェックといった AI の外側の人間の行為に責任原因を求める構造であるという点を明確にしたほうがよい。

- 判断の支援にとどまる補助 AI か、判断を全面的に委ねる代替 AI かを明確にした方がよい。
- 事例内での具体的な損害等が明確ではないため、不利益の内容や事実関係の具体的な記述が必要である。
- 取引審査 AI による不利益な取扱いに対して個人が疑念を持っただけでは、社会問題として顕在化しにくい。より現実的に即した事例とするなら、内部告発や SNS での告発など、問題が発覚し得る経路も考慮することが重要である。
- 資料 14 頁について、運用開始後に差別的な挙動がないかをチェックする事後モニタリングが実務的に重要であり、こうした措置が法的に評価されるインセンティブ設計が必要である。その点についても、記載いただけるとよい。
- AI による差別は統計的な傾向として顕在化するため、個々の出力に対して差別回避の注意義務を課すことは現実的ではない。したがって、単に注意義務を負わせるのではなく、開発者と利用者が協力してシステムを改善していくインセンティブ設計が重要である。
- 開発段階でデータセットの不備等によりシステムティックに差別が生じる場合と、運用を通じて事後的にバイアスが判明する場合を区別して、議論の解像度を高める必要がある。

#### 【AI 利用者の責任について】

- 事務局資料が、AI の判断内容がブラックボックスであることを理由に利用者が過失責任を免れるという趣旨だとすれば、疑問がある。解釈論なのか立法論なのかはさておき、ブラックボックスな判定結果をそのまま利用するのであれば、その内部で起こっていることも含めて責任を引き受けるべきという評価も考えられる。
- 本事例における AI を補助 AI と位置付けるならば、議論は利用者の判断の適切性に収斂する。その場合、因果関係は「注意義務違反（過失）」と「権利侵害」の関係として整理されるため、AI のバイアスに起因する因果関係という特殊な問題を独自に論じる余地は大きくないと考えられる。

#### 【因果関係について】

- 取引拒絶の事案は、原因特定が困難な場合があり、水俣病訴訟などで用いられた確率に応じた損害額の認定は難しいのではないかと。また、生命・身体とは被侵害法益が異なるため、相当程度の可能性論や期待権論をそのまま適用することにも疑問が残る。

- 単なる証拠の欠如とは別に、AIの性質に照らし因果関係が確率的にしか論じられないという事案が生じた場合には、確率的な挙動をするという対象の性質に着目して割合的に因果関係を評価することもあり得るとは思われる。
- 実務的には、データセットの偏りなどから差別的な結果が生じたことは経験則上推認可能であり、事実上の推定が活用され、開発者側に反証が求められる可能性がある。
- システマティックに特定の属性が不利益を被るようにアルゴリズムが組み込まれている場合は、確率的な挙動であったとしても、因果関係が認められやすいと考えられる。意図的に差別した場合と結果的にそうなった場合は分けて考えるべきである。
- 代替法益としての「相当程度の可能性」論は、通常は生命・身体的な法益に関連して論じられるため、本事例のような契約締結の可能性という利益に適用することには疑問が残る。期待権論についても、契約交渉の不当破棄に関する裁判例を参照・検討するほうが適切ではないか。

#### 【代理変数と差別について】

- 差別があればすべて注意義務違反になるということではなく、社会的相当性や合理性がない差別によって取引を拒絶することが問題であるということをも明記したほうがよい。
- 資料13頁の「代理変数」の表現は、何が何の代理変数なのか分かりにくい。直接的な差別を意図していなくても、収入や勤務形態といった変数が性別などのセンシティブ属性と社会構造的に結びつき、結果として間接差別につながる可能性があるため、誤解を招かないよう慎重な記述が求められる。
- 私人間における別異取扱いが直ちに違法となるわけではないが、社会的許容限度を超える場合には問題となり得る。ただし、公法上の規律が存在する業種・業態においては、その点を考慮して、私法上の不法行為の成否等が判断されることになる。
- 本事例のように、多数の不動産事業者がAI開発者による同一のプラットフォームを利用している場合、特定の個人が多くの人から差別・排除されるリスクが生じる。この場合、被害者がAI開発者を直接訴えることは困難であるため、公法的な規律を含めた解決策の検討が必要である。
- AI活用の際の人間の関与度合いとイノベーションへの配慮について、業務や利用形態ごとにリスクの程度は異なるため、一律に対策を求めるとリスクとイノベーションのバランスを損なう恐れがあり、事例ごとに慎重に検討する必要がある。

#### 【資料中の表現について】

- 資料の記載において、「権利侵害」の有無と「過失」の有無が混在して論じられており、報告書の取りまとめに際して留意されたい。
- 資料7頁の「属人的バイアス」については、「認知バイアス」の趣旨で用いているので

あれば、その旨を明確にし、「統計的バイアス」と対比させるべきである。

- 資料 11 頁の「ルールベースの審査を行っていない」という記述は、全ての AI に当てはまるわけではないため、「機械学習している AI については」といった限定的な表現とするほうが的確である。

### 3) 想定事例 6 について

資料 3 の 20～37 頁に基づき、事務局から、想定事例 6（自律走行ロボットの事例）について説明が行われた後、構成員による討議が行われた。

#### 【事例の位置づけについて】

- 本事例は、人間の作業を代替するだけでなく、人間にはできなかった作業も行うという「プラスアルファの効用」が期待される機械である点、また利用者が事業者であり一定の注意義務を要求し得るといった点で、自動運転車とは特徴が異なる。このように、製品の特性（プラスアルファの効用の有無）や利用者の属性（事業者か消費者か）によって議論の射程が異なることを、報告書の冒頭で明確に示すべきである。
- 利用者が事業者である本事例では、自動運転車で議論されるような、ユーザーの過度な期待や依存傾向を前提とした欠陥判断は、基本的に当てはまらないか、当てはまる度合いが小さいことを注記すべきである。

#### 【安全性の判断基準について】

- 資料 31 頁の整理等は実質的に人間ではなく AI と比較しているような記述となっている。したがって、判断基準の整理としては、「同種の AI」との比較は困難であるとの批判があることは踏まえつつも、「人の能力」だけでなく、「別の AI システムの安全性」を基準とする考え方も明確に位置づけて整理する必要がある。
- 資料 26 頁の整理について、自動運転の文脈では、ルールベースを前提とした個別状況の安全性を基準とする考え方でなく、機械学習の進展に伴い、走行距離あたりの事故率といった統計的な考え方も有力になっている。

#### 【欠陥の種類と責任について】

- 資料 27 頁の欠陥種類の整理について、ソフトウェアのバグが「製造上の欠陥」に分類されている点に違和感がある。AI の検知失敗が「設計上の欠陥」とされていることと合わせ、欠陥の分類が一貫していないように見受けられる。
- 自動運転車に比べ、指示・警告上の欠陥に関する記述が少ない。利用者が事業者であっても、指示・警告が複雑で理解困難な場合は、システムの複雑性自体が設計上の欠陥と評価される可能性があり、その点を盛り込むべきである。

- 資料 30 頁について、本事例は自走式だが、もし工場等の床に固定されたアーム型ロボットであれば、土地工作物責任（民法 717 条）の適用対象となり得るため、立法論としては瑕疵を要件とする無過失責任を課すことも考えられる。
- 資料 31 頁の「予測困難な挙動」という表現について、AI の出力が原理的に予測できないことは自明であるため、この言葉だけでは概念として広すぎる懸念がある。AI の出力が予想できないという一般的な性質と、それを超える高度な予測困難性とを区別して記述すべきである。
- 資料 31 頁で開発危険の抗弁の可能性に言及しているが、現行法上、製造物責任法における開発危険の抗弁の適用は極めて限定的である。また、イレッサ訴訟最高裁判決で示された予見可能性の考慮も、製造物一般に直ちに適用できるものではないため、こうした考え方もあり得るといった程度の位置づけに留めるべきである。
- AI は確率的に挙動するため、長期的・統計的に見て安全性が担保されているのであれば、たまたま人間と異なる挙動で事故を起こしたとしても、直ちに「欠陥」と認定すべきではない。稀な挙動の全てに責任を負わせると開発の萎縮を招くため、個別の事故回避能力ではなく、リスクを生じる具体的場面に即して、統計的な安全性を重視する運用上の工夫が必要である。

#### 【自己診断 AI の製造物性について】

- 資料 34 頁について、ロボット本体の画面に表示される形式であれば、物理的な製造物の枠内に収まるため、誤った指示が表示された場合は、動作上の欠陥、あるいは指示・警告上の欠陥として製造物責任を問うことができる。また、独立したホスト PC 等で診断結果が表示される場合でも、製造業者が AI を利用してユーザーに指示を与えている以上、指示内容が不合理であれば、その責任は免れないと考えられる。
- 自己診断機能を提供する LLM が物理的に離れたサーバー上にある場合でも、ロボット本体がその機能の利用を予定しているのであれば、利用者は機能全体の安全性を期待するため、当該機能も製造物の一部と捉えることで差し支えない。

#### 【アップデートと欠陥判断基準時について】

- 製品の引渡し後にアップデートが行われることが前提となっている製品においては、メーカーが管理権限を持ちアップデートを行っている以上、実質的には製品の内部に対する支配が継続していると捉えられる。したがって、リコールで回収して基盤を交換する場合と同様に、内部的な引渡しは完了していないと構成し、アップデート時点での欠陥を問うことが可能ではないか。
- 資料 35 頁について、現行法は「引渡し時」を欠陥判断の基準としており、単に「アップデートが予定されている」という一般的な可能性のみをもって引渡し時の欠陥とす

るのは解釈上困難である。欠陥が肯定されるのは、アップデートによる不具合の潜在的な原因が引渡し時に既に存在していた場合に限られると考えられる。

- 引渡し後にバグが発見されたにもかかわらず、メーカーがアップデートを提供しなかったことによる不具合（不作為）も検討に値する。現行法での対応は困難な面もあるが、製品の性質上、本来カバーすべき責任の空白が生じていると考えられる。
- 「予測困難な挙動」と開発危険の抗弁の関係について、アップデートを行わなかったことによる責任について、被害者に過失の立証を求めることは困難であるため、開発危険の抗弁をインセンティブ設計として活用できるように、「引き渡した時」の解釈を工夫すべきである。すなわち、事業者が組織的なモニタリングとアップデートを行っていた場合に限り、当該製品を利用に供した当時の知見では回避不能だったとして抗弁を認める運用とすれば、開発者は免責を得るために積極的に製品改善を行う動機付けが働くことになる。
- 資料 35 頁について、以前「潜在的なバグがあれば引渡し時の欠陥と見てよい」と述べたが、それは現行法の「引渡し時」という要件に引きつけて考えざるを得ないためである。本来的な形としては、「引渡し時」という言葉の意味を AI との関連で明確にするための解釈が必要である。

#### 【立証責任（事実上の推定）について】

- 資料 36 頁について、従来の「欠陥の事実上の推定」に関する判例法理は、人間が対象物をコントロールしていることを前提に、通常の使用で異常が発生したから欠陥があるはずとするものである。しかし、AI 搭載機器は自律的に作動し、人間によるコントロールが後退するため、「動作の異常性」の判定自体が困難となる。従来の法理は事故の科学的な機序の立証負担を軽減する意義があったが、AI の文脈でその機能が発揮されるかは疑問である。本来であれば、技術的・科学的な複雑性に正面から対応する法理を含めた検討が必要である。

#### 4) 全体を通して（座長総括）

##### 【想定事例 5 について】

- 従来の事例とは異なり、判断構造そのものまで踏み込んで検討する必要がある点や、「代理変数」等の記述に関して注意が必要であるとの指摘に留意したい。
- この事例は判断補助 AI であるとの意見があったが、そのように理解できる。
- 因果関係について、確率的な枠組みで考える必要性はあり得るものの、実務的には事実上の推定で対応可能なケースも多いとの指摘もあった。他方で、生命・身体侵害の事案とは異なるため、「相当程度の可能性」の法理をそのまま適用することは難しく、確率的に考える場合の理論構成については、引き続き検討が必要である。

- 利用者による事後モニタリングが極めて重要であり、開発者と利用者が共同して問題を解決していく方向性が重要であるとの意見は重要である。
- また、業種・業態ごとの公法的な規律の存在も重要な考慮要素となるほか、AI 利用における人間の関与度合いやリスクも業種や業務ごとに異なるため、報告書の作成においては、これらの点に配慮が必要との指摘もあった。

#### 【想定事例 6 について】

- 利用者が消費者ではなく「事業者」である点は重要なポイントであり、指示・警告上の欠陥についても記述を充実させる必要がある。
- 安全性判断において、自動運転等の分野では統計的な視点が有力になりつつある点や、比較対象として人間だけでなく他の AI システムを基準とする考え方がある点についても、報告書に反映したい。また、固定型ロボット等との関係で、土地工作物責任（民法 717 条）の適用の可能性についても指摘があった。
- アップデートに伴う欠陥については、意見が分かれたものの、メーカーがアップデートを行った以上は製造物責任を負うという考え方に賛同する。もっとも、現行法解釈との整合性について慎重な意見もあるため、引き続き検討が必要である。
- アップデートしなかったことによる不具合については、従来は過失責任の範疇と考えられてきた。なお、欠陥として構成する可能性や、開発危険の抗弁との関係について、引き続き検討したい。

## 5.4 第4回

### ① 開催日時

令和8年1月14日（水）15:00～17:00（オンライン会議）

### ② 議事次第

1. 開会
2. 議事
  - (1) 第3回研究会の議論の内容について
  - (2) 報告書案に関する検討
3. 閉会

### ③ 配布資料

- 【資料1】 議事次第・配布資料一覧
- 【資料2】 構成員等名簿
- 【資料3】 事務局説明資料
- 【資料4】 AI利活用における民事責任の解釈適用に関する手引き（案）（非公開）

### ④ 出席者

大塚座長、稲谷構成員、江間構成員、柿沼構成員、宍戸構成員、白石構成員、中原構成員、橋本構成員、福岡構成員、松尾構成員、オブザーバー（関係省庁・関係団体）、事務局

### ⑤ 議事概要

#### 1) 第3回研究会の議論の内容について

資料3の3～4頁に基づき、事務局から前回研究会の討議概要と対応方針について説明が行われた。

#### 2) 報告書案について

資料3の5～10頁および資料4に基づき、事務局から本研究会の報告書案に関する説明が行われた後、以下のような討議が行われた。

#### 【AIの類型（論点1）：事例区分の名称について】

- AIの利用形態を、2つに区分して整理する方向性については、賛成である。
- AIの類型の名称について、AIの判断を尊重しながら用いるという点は、「補助利用型」でも同様である。AIの自律的判断を予定し、これを受容することに本質があるとする、「判断補助利用型」や「自律判断受容型」といった名称も考えられるのではないかと。

- 「判断尊重型」という名称は、AIの判断を尊重する必要性や尊重する義務があるかのような誤解を招く懸念がある。また、判断尊重型 AI に関する記載には、必ずしも判断を尊重しない場面が含まれると考えられることとも整合しない。「補助型」、「自律型」といった簡潔な名称としたほうが良いのではないか。
- 「判断尊重型」という名称について、判断や行動を代替する前提で提供されるといった記載があるところ、「代替」ではなく「尊重」とした理由が分かりにくい。また、画像生成等の文脈では「判断」よりも「出力」と表現するほうが正確である。名称の主語が異なったりすると誤解が生ずるため、「判断補助型」と「判断代替型」、あるいは「補助型」と「代替型」といった名称が良いのではないか。
- 「判断尊重型」という名称について、やや分かりにくい印象を受ける。「判断補助型」と対比させたときに対象が不明確であり、「尊重」という言葉にも違和感がある。「代替」や「依存」といった案もあるが、一般的な読者にとっての理解のしやすさや、英訳されることも踏まえた用語選定が必要である。
- 「判断尊重型」という表現は、AIが何かの判断を尊重しているかのような誤解を与える懸念がある。本来的には「AI出力重視型利用」などが良いかもしれず、再考の余地がある。
- 「補助」と対になる概念として、「受容」という表現を用いる案もあるが、「補助」として利用する場合も受容する側面があるため、厳密には対比が正確ではない。AIの挙動を監督して使う「補助」に対し、AIの判断に従って処理するという意味での「依拠」も考えられ、「補助」と対になる概念を上手に使った方が良い。

#### 【AIの類型（論点1）：事例の整理や説明について】

- 各類型内部の多様性について整理が必要である。想定事例6（自動走行ロボット）について、製造物責任から論じるという順序は、AIを組み込んだ製造物の安全性が本質であるため正当であるが、このような構造の特徴について、冒頭で言及すべきである。「判断尊重型 AI」の事例も多様であり、ある程度の偏差があることを説明する必要がある。また、「補助利用型 AI」について、想定事例1（配送ルート最適化 AI）は人の判断を AI が代替しているとはいえないケースである。想定事例2（弁護士業務支援 AI）、3（画像生成 AI）、4（取引審査 AI）は、AIが人の判断を代替する点で共通するが、想定事例2は、最終的な判断が人でなければならないケース、想定事例3、4は AI の判断について一定の検証が求められるケースと整理できる。その上で、想定事例3、4は、AIの出力自体が権利侵害を内包し得るため、利用者がそのまま出力結果を使ってはならないという点で共通するが、出力の権利侵害性や対処の必要性をどれだけ人が容易に判断できるかという点に違いがあり、利用者が負う義務の在り方が異なる。

- 人の操作や介入を前提としないシステムに関する裁判例に関する記載は、各類型に直接的に関係しないため、記載場所を移動すべきである。また、判断尊重型 AI に該当するための必要性について、作業効率の顕著な向上だけでよいのかは再考すべきである。さらに、補助利用型 AI 及び判断尊重型 AI の区分に関する記載の構成について、まず「補助利用型 AI」にしか該当しない場合を整理した上で、どちらの類型にも当たり得る場合について判断補助型 AI に該当するための要件を説明したほうが分かりやすい。

### 【AI エージェント（論点 2）について】

- AI エージェントに関し、利用体制の構築及び運用に関する注意義務の記述については、全体としては、違和感はない。
- もっとも、AI エージェントが外部ツールやシステムと連携してアクションを起こす際のリスクの観点からの言及を追加することが考えられる。
- AI エージェントの想定事例について、「判断尊重型 AI」に当たり得るものの一例とするのであれば、AI エージェントが担う業務をより高度な、AI に判断を委ねる必要性がある事例とすることが望ましい。現状の想定事例は分かりやすいが、「判断尊重型 AI」に該当する必要性が高い印象は受けなかった。
- AI エージェントの回答が取引の勧誘に当たることを前提として付記すべきと思われる。また、判断尊重型 AI に当たる場合に、AI 利用者がいくら注意を尽くしても AI が誤った回答をすることがあるが、このような場合であっても AI 利用者に責任を負わせる余地はあると考える。

### 【想定事例 3（画像生成 AI）について】

- 想定事例 3 の事例 a について、AI の機能自体は画像生成を行うものであり、パブリシティ権侵害になるかどうかの判断は人が行うこととなり、AI の機能や射程と異なる。したがって、「AI の判断を尊重する」といった表現は馴染まないのではないかと。
- 想定事例 3 の事例 b における幫助責任について、カラオケ装置のリースに関する裁判例が挙げられているが、カラオケの場合は利用店へのリースという形態上、権利侵害の蓋然性が高く、相手方を選ぶ義務が生じるのに対し、AI は一般的に提供されるものであるため、権利侵害の可能性を下げる措置等が注意義務の中心になるという違いがあるものと思われる。
- 想定事例 3 の事例 c について、他の事例との連続性が薄いため、記述を分ける等の工夫をした方がよいと思われる。
- 想定事例 3 の記載について、AI 利用者の責任と AI 開発者・提供者の責任が混在しており分かりにくいいため、見出しを追加するなどして責任の主体ごとに整理すべきである。

る。

- 正則化項に関するコラムの記述について、やや簡略化しすぎている面はあるものの、技術的な内容を簡易に伝える内容としては妥当と考える。

#### 【想定事例 4（取引審査 AI）について】

- AI 利用者の責任について、AI の判断過程に対して AI 利用者が注意義務を負うとの記載になっているが、これが「判断尊重型」の責任に近い記載となっているため、整理が必要である。
  - 想定事例 4 について、出力結果のバイアスを一見して判断することは困難であるため、AI 利用者の最終的な判断において適正な範囲で AI の出力を利用するためには、モデル選定等の判断過程の注意義務が主軸になると整理した。これは、AI の特性に応じた注意義務と考えている。（事務局）
  - 想定事例 4 については、【利用者が】最終的なアウトプットをコントロールする責任はあるものの、AI による個々の出力の当否を利用者が判断し難いことが特徴である。そのため、アウトプットの適切性をコントロールする責任として、【利用者が】データを集めて介入できる仕組みを作っておくという注意義務にならざるを得ない。判断尊重型 AI と注意義務が類似する面はあるものの、このように明確に区分し難い事例が出現し得ることも AI の特性の一つと考えられる。
  - 想定事例 4 は、個々の出力の検証が困難である点において「判断尊重型」にも当たり得るグレーゾーンの事例であるため、各事例の性質の違いを明確に整理すべきである。
- 想定事例 4 について、貸貸業者が AI の出力を採用している以上、その出力については貸貸業者が責任を負うと端的に考えるべきではないか。
- 取引審査 AI に関するバイアスや品質モニタリングについては、一般的な手法が定まっているわけではないため、具体的な記載を検討するのが難しい。

#### 【想定事例 5（外観検査 AI）について】

- AI 提供者・利用者の責任についての記載箇所に判断尊重型 AI への該当性に関する記載があるため、構成上、別の箇所に移動すべきである。

#### 【想定事例 6（自動走行ロボット）について】

- 欠陥判断の方法について、自律的に動作する機械に即した統計的基準や個別状況基準がうまくいかないから従来型基準を用いる、という論理構成は適切か。AI のリスクというよりは機械の危険性が問題となる事例も多いため、従来型の欠陥判断基準で対応できることを明確に記述した方がよいと思われる。

- 現状では、従来の工業機械等と同様の考え方で欠陥を判断するとされている論理構成について、自律的に動作する機械特有の要素と従来の要素とを総合的に考慮するといった整理が適切ではないか。

#### 【全体の構成・表現について】

- 一般読者にとっての分かりやすさを向上させるため、6つの想定事例の関係性や全体像を示す図表を作成することも考えられる。
- 法律の専門家ではない一般の読者にとっても、どこを見れば大枠がつかめるかが分かるよう、概要の作成や図示等の工夫が必要である。
- 本報告書で取り上げた事例は、網羅的なものではないことを明記したほうがよい。また、技術の進展に伴って将来的なアップデートがあり得るのであれば、その点も明記したほうがよい。
- AI開発者・提供者の責任について、報告書全体として「例外的」と「限定的」という用語が混在しているため、「限定的」に統一すべきである。
- AI開発者・提供者が「補助利用型AI」として提供しているが、AI利用者が「判断尊重型AI」として使ってしまう場合など、認識の齟齬により問題が生じるケースも想定されるため、AI開発者・提供者による説明責任が重要になる点に留意すべきである。

#### 【資料4・総論部分について】

- 総論部分の不法行為責任が当事者間の責任を形成するデフォルト・ルールであるという趣旨の記載について、当事者間の契約による責任関係の明確化においても、対被害者の損害賠償請求が重要であることをより明確に記載すべきである。
- 総論部分で引用されているATMの預金払戻しに関する裁判例について、当該裁判例が「判断尊重型AI」に直結するかのような記述は誤解を招くため、あくまで当該裁判例を押し進めるとどうなるかといった表現に修正すべきである。
- 判断尊重型AIの注意義務として、「AIシステムの利用体制の構築」等の記載があるが、「利用」という言葉が狭く捉えられる懸念があるため、「AIシステムを組み入れた業務プロセスの構築」等の表現も検討すべきである。

### 3) 全体を通して（座長総括）

- 構成員の皆様の議論に改めて感謝申し上げたい。今後の報告書案の最終的な検討・確定については、座長一任とさせていただきたい。（異議なし）
- 多くのご意見をいただいた「判断尊重型」の類型の名称については、「出力依拠型」や「出力受容型」等の他の候補も踏まえて、事務局と改めて検討したい。

#### 4) 経済産業省挨拶

- 全4回の研究会における精力的な討議に深く感謝申し上げたい。本研究会での議論は、今後各業界におけるガイドラインや契約等において、個別具体的な責任分界を明らかにしていく際にも重要な基礎となるものである。政策セクションとしても、議論をしっかりと消化し、世界で最も AI を開発・活用しやすい国の実現に向け、政策反映を図っていききたい。

令和7年度技術開発調査等推進事業  
(AI利活用における民事責任の在り方に関する調査研究及び研究会運営)

調査報告書

令和8年3月

みずほリサーチ&テクノロジーズ株式会社

デジタルコンサルティング部

〒100-8176 東京都千代田区大手町一丁目 5-5