

# 第3回 AI利活用における民事責任の在り方に関する研究会 事務局説明資料

2025年12月3日

商務情報政策局 情報経済課

1. 第2回研究会でのご意見
2. 想定事例5：取引審査AI
3. 想定事例6：自律走行ロボット（AMR） - 前半
4. 想定事例6：自律走行ロボット（AMR） - 後半

# 第2回研究会の討議概要①

区分	ご意見の概要	事務局対応方針
想定事例3について	<ul style="list-style-type: none"><li>● AIサービス開発事業者A、検品受託者Bの注意義務について、基本的に異存はない。</li><li>● 過失判断の対象が、「人間による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」へと質的に転換するとの理解が広く支持されたが、人が判断を行う場合と比べて「注意義務が軽減した」ということではないため、留意を要する。</li><li>● 過失判断の対象が転換する前提として、検品についてAIシステムを用いること自体の適否の問題があり、検査の精度が十分に高いか否か等は、この適否を判断する基準と捉えることができる。</li><li>● 過失判断の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ることができる点は明確にしておいた方がよい。</li></ul>	<ul style="list-style-type: none"><li>● 左記を含め、いただいたご指摘については取りまとめに反映させていただく。</li></ul>
専門家の注意義務について	<ul style="list-style-type: none"><li>● 認知科学の知見からは、専門家であってもAIによる注意水準への影響は避けられない。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判断能力」こそ、専門家としての注意義務の中心となるべきではないか。</li><li>● 他方、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ずるとしても、弁護士等の専門家においては、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。</li></ul>	

## 第2回研究会の討議概要②

区分	ご意見の概要	事務局対応方針
想定事例1から3の整理について (資料14頁)	<ul style="list-style-type: none"><li>想定事例1から3の整理について、従来の過失判断に沿った整理が可能である、A)のAI開発者の義務はB)でも妥当し、B)のAI開発者の義務がAI利用者との合意で規定されるという内容はA)でも妥当する、資料14頁右側の4つの考慮要素は場面に応じて使い分けることが必要である、等の指摘がなされた。</li></ul>	<ul style="list-style-type: none"><li>左記を含め、いただいたご指摘については取りまとめに反映させていただく。</li></ul>
想定事例4について	<ul style="list-style-type: none"><li>想定事例4の整理について、概ね異存はない。</li><li>参考裁判例⑥との関係で故意・過失に関する記載を補強する必要がある、参考裁判例⑩との関係で想定事例4を同様に整理して良いかは検討が必要である等の指摘がなされた。</li></ul>	
立証上の論点について	<ul style="list-style-type: none"><li>文書提出命令の相手方が訴訟当事者の場合と第三者の場合で、要件や運用に差異が生ずるかという点は論点となり得る。</li><li>証拠を出すより、敗訴や和解によって金銭解決によるほうが得だから証拠を出さないとなると、情報が集まらなくなるため、民事訴訟以外の制度との連携も視野に入れて、社会全体でAIの安全性を確保する仕組みの設計が重要である。</li></ul>	
因果関係について	<ul style="list-style-type: none"><li>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。</li></ul>	<ul style="list-style-type: none"><li>第3回以降の議論では、いただいたご指摘も踏まえて検討してまいりたい。</li></ul>

上記のほか、第2回研究会の論点について様々なご意見を頂いた。

1. 第2回研究会でのご意見
2. 想定事例5：取引審査AI
3. 想定事例6：自律走行ロボット（AMR） - 前半
4. 想定事例6：自律走行ロボット（AMR） - 後半

## 想定事例5：取引審査AI

- AIサービス開発事業者A社は、賃貸物件の入居申込者の信用スコアを算出する入居審査AIを開発・提供している。当該AIサービスでは、利用者である不動産業者が入居申込者に関する所定の情報をインプットすると、当該情報に基づき信用スコアを定量的に算出した上で、①合格、②最終的な合否を人が判断する、③不合格の3ランクに分類する。Aは当該AIサービスをクラウド形式で複数の不動産事業者に対し提供している。
- 不動産業者B社は、AのAIサービスを導入し審査業務を効率化・省人化していた。入居申込者Vは、Bが保有し賃貸に供している物件の一つへの居住を希望し、当該物件への申込みを行った。AIはVの信用スコアが上記②のカテゴリに該当すると判断したため、Bの担当者B'はAIの出力を参考にしつつも、改めてVの信用情報を自社基準と照らし合わせて判断し、Vの申込みを却下した。
- 後日、Vは同じ職場の同僚が、Vが希望していたものと同価格帯のBの賃貸物件に応募し、AIによる審査を受けたところ、AIが合格と判断し希望の物件に入居したとの話を聞いた。当該同僚は資産や収入、雇用形態等において似た境遇であったため、VはAIに性別などの属性に着目した不公正なバイアスが含まれており、そのために低評価を受けたのではないかとの疑念を持った。



# 入居審査AIにおける統計的バイアス

- 入居審査を含む取引審査には一般に取引自由の原則が妥当し、契約相手方の選定や取引条件の設定について裁量が生ずるが、AIの出力が差別的な取引拒絶に繋がった場合には、A及び／又はBの不法行為を構成する可能性がある。
- 入居審査AIは個人に関する様々な属性を考慮して下記のような分析を行う。
  - AIの活用により、多様な要素を考慮した精度の高いリスク分析が可能となることが期待され、また担当者による属人的なバイアスを低減することも期待される。
- ```
graph LR; A["
  - 本人の属性情報：氏名、年齢、性別、居住地 等
  - 経済的情報：収入、勤務先・勤続年数、資産状況 等
  - 同居者の情報：同居者数、関係性、年齢・職業・収入 等
  - 物件情報：物件所在地、家賃、間取り 等"] --> B["
  - 家賃滞納リスク
  - 滞納発生時期
  - 滞納時の損失 等"]; B --> C["
  - スコアリング
  - ランク評価 (A～C)"]
```

  - 本人の属性情報：氏名、年齢、性別、居住地 等
  - 経済的情報：収入、勤務先・勤続年数、資産状況 等
  - 同居者の情報：同居者数、関係性、年齢・職業・収入 等
  - 物件情報：物件所在地、家賃、間取り 等
  - 家賃滞納リスク
  - 滞納発生時期
  - 滞納時の損失 等
  - スコアリング
  - ランク評価 (A～C)
- 入居審査を含む取引審査には一般に取引自由の原則が妥当し、契約相手方の選定や取引条件の設定について裁量が生ずるが、憲法14条1項後段に列挙される属性等の差別に繋がりうる属性（以下「センシティブ属性」という）<sup>1</sup>を理由とした不公正な取引拒絶と評価されるケースでは、人格権侵害に基づく不法行為<sup>2</sup>に該当する可能性がある（後記9-10頁・参考裁判例①～⑥参照）。
- 日本銀行金融研究所『「金融機関におけるAIの利用を巡る法律問題研究会」報告書—金融機関におけるAI利用に伴う私法上のリスクと管理』21頁（2025年6月）においても、「例えば、自律型のAIを用いた融資審査において、AIの判断に不公正なバイアスが反映されていた場合には、融資審査を公正に行うという信義則上の義務に違反しうる。」と指摘されている。

1 個人の特徴のうち、どこまでの範囲について公平性が要求されるかは、それぞれの国や地域における文化的・社会的背景によっても異なり、一義的に確定することは困難である。そのため想定事例5では、従来の裁判例で不法行為の成否が論点となってきた性別や国籍に基づく契約拒絶を想定し、責任論上どのような点がポイントとなるかについて検討を行う。

2 取引に付随する信義則上の義務違反に基づく不法行為とされる場合もある（後記10頁・参考裁判例⑥参照）。

# 従来の取引拒絶型の裁判例における判断枠組み

- 従来の裁判例では、以下のような要素を相関的に考慮している。
- 結論として違法と評価されたのは、主にセンシティブ属性を直接的に取引拒絶の根拠とした場合であり<sup>3</sup>、このような場合、目的の合理性や手段の適切性が特に求められる。  
→AIの利用過程において、このような権利侵害を生じないようにするため、AI開発者・提供者・利用者がそれぞれどのような注意義務を負うか。

## (1) 当該取引における相手方の属性の重要度

- 相手方の属性が一般に重視されず広く開放されていたり、公共性が認められる財やサービスについての契約では、契約締結等の自由にも一定の限界があり、拒絶理由の合理性等が求められやすい（後記9-10頁・参考裁判例③、④）。他方、相手方の属性が重要な取引では、契約の相手方や内容の決定について広い裁量が認められる傾向にある（後記9-10頁・参考裁判例②、⑤）。

## (2) 不利益の程度

- 取引の機会が失われることの不利益の大きさ等が問われ、取引拒絶の対象となる者に代替サービスが多く提供されている場合など、取引拒絶による不利益が小さい場合には不法行為が成立する可能性は低減する（後記9-10頁・参考裁判例①、⑤）<sup>4</sup>。

## (3) 目的の合理性

- 上記(2)における不利益の大きさ等に応じ、取引拒絶等を行う目的の合理性が求められる（後記9-10頁・参考裁判例①～④）。
- 取引の性質に鑑みて取引条件を定形化する必要性が高い等の事情がある場合には、センシティブ属性に基づく取引条件の付加について目的の合理性を認めた事例がある（後記9頁・参考裁判例②）。

## (4) 手段の適切性・代替手段の有無

- 目的との関係で一律に取引を拒絶することの合理性や、より間接的・具体的な属性に着目した代替手段の有無等が問われる（後記9頁・参考裁判例③）。

3 本資料で引用したほか、東京地判平成16年9月16日（判例集未登載）、東京地判平成24年11月2日（判例集未登載）、大阪地判平成29年8月25日判時2368号23頁等も、差別が禁止される属性を直接的な根拠とする取引拒絶が違法と判断された事例である。

4 取引拒絶自体の経済的不利益が必ずしも大きくない場合であっても、人格的利益の侵害の重大性を踏まえて違法性が認められている事案も存在する（後記10頁・参考裁判例④）。

## 参考裁判例（1/2）

### 参考裁判例①：女性専用車両の提供（東京地判平成23年7月12日（判例集未登載））

- 女性専用車両の提供は、平日の通勤通学の時間帯に相当な混雑をする首都圏等大都市圏の通勤電車において、痴漢犯罪の被害を受けるおそれのある女性の乗客に対し、少しでも安心、快適な通勤通学環境等を提供するために行われていると解せられ、これは目的において正当というべき。
- 女性専用車両が設定されるのは、平日の通勤時間帯の一部電車で、6両編成の車両のうちわずか1両のみに過ぎず、これは健常な成人男性の乗客をして他の車両を利用して目的地まで乗車することを困難ならしめるものではないから、健常な成人男性の乗客に対し格別の不利益を与えるものでもない。

➤ 目的の合理性及び手段の適切性（特に、車両の時間帯や数が限定的に運用されていること）を理由に不法行為の成立を否定した。

### 参考裁判例②：住宅ローン審査における永住資格要件（東京地判平成13年11月12日判時1789号96頁、東京高判平成14年8月29日金商1155号20頁）

- 銀行が永住者の在留資格のない外国人からの住宅ローンの申込みを拒絶した行為は、憲法14条などに違反するものではない。
- 住宅ローンは、長期の返済期間を設定すること等から融資に要するコスト（契約事務費、債権・担保管理費、債権回収費用など）が相対的に高く、採算性の高い商品であるとはいえないから、銀行が十分な採算を取るためにには大量の住宅ローンを取り扱うことが必要であり、そのためには、貸付対象者等の要件をできるだけ定型化する必要がある。
- 永住資格を有しない外国人は在留期間が3年間に制限されており、本邦を出国した場合、その債権管理及び債権回収に要する費用、時間、労力が多大なものにならざるを得ないこと等を踏まえると、在留資格の有無によって定型的に融資資格を判断することは合理性に欠けるものではない。

➤ 住宅ローンという契約の相手方の属性が重要な取引において、銀行の営業の自由を踏まえつつ、定型的な取引条件を設ける必要性や、借主の出国による債権回収費用の増大等を考慮し、一律に永住資格要件を求めることが合理性を認定した。

### 参考裁判例③：公衆浴場における入浴拒否（札幌地判平成14年11月11日判時1806号84頁）

- 被告の施設は公衆浴場法に基づく知事の許可を受けて経営されており、公共性が認められることを踏まえ、すべての外国人の利用を一律に拒否するのは明らかに合理性を欠くものというべきであり、他の入浴者に迷惑をかけたり入浴マナーを守らない者に対しては個別に指導・対応すべきであったと判示して、不法行為の成立を認めた。
- 被告施設の公共性を踏まえ、国籍に基づく一律の入浴拒絶ではなく、個別に指導・対応を行う等の代替措置を行うべきであったと指摘して不法行為の成立を認めた。

## 参考裁判例（2/2）

参考裁判例④：ゴルフクラブの入会拒絶(1)（静岡地浜松支判平成26年9月8日判時2243号67頁）

- 女性への性別変更の審判を受けた原告に対する入会拒絶の事案において、被告のゴルフクラブには厳格な入会要件が存在しないこと等から、団体としての閉鎖性が認められないとした上で、原告が被告のゴルフ場を用いることによる既存会員への強い不安感や困惑といった危惧は抽象的なものにすぎず、他方、原告が被る精神的損害は看過できないものであるから、被告による入会拒絶は違法と判断された。

参考裁判例⑤：ゴルフクラブの入会拒絶(2)（東京高判平成14年1月23日判時1773号34頁・原審東京地判平成13年5月31日（判例集未登載））

- 参考裁判例④とは異なり、入会手続の厳格さ等に鑑みて閉鎖的な性質が高いと認められた事案では、外国人の入会を一律に制限している場合であっても不法行為の成立が否定された。
- 原告の不利益の小ささを示す事情として、会員資格を取得したとしても、ゴルフ場の利用等における便宜が得られるに過ぎないことが指摘された。

- 団体としての閉鎖性が認められなかったケースでは、性別変更を考慮することの合理性や原告に対する不利益の程度を考慮した上で、不法行為の成立が認められている。
- 他方、閉鎖性が高いと認められたケースでは、一律に国籍要件を定めるケースでも不法行為の成立が否定されている。

参考裁判例⑥：外国籍であることを理由とする賃貸物件への入居申込みの却下（京都地判平成19年10月2日（判例集未登載））

- 賃貸マンションの所有者である被告が、原告が入居することを前提として契約書を作成し、契約を締結する段階に至っていたにもかかわらず契約を拒絶した事案において、原告が日本国籍でないことは契約拒絶の合理的理由に当たらないと判示し、被告の不法行為責任を認めた。
- 損害額については、原告の主張する逸失利益等の損害が発生したとは認められなかつたが、慰謝料及び弁護士費用として110万円の賠償が命じられた。

- 契約締結段階における契約成立への信頼の保護に関するものであり、不公正な取引拒絶に関する判断を示したものではないが、入居審査の文脈において、原告が日本国籍でないことは拒絶の合理的理由に当たないと判示した。

# 審査にAIを用いる場合の特殊性

- 契約締結拒絶に関する従来の判断枠組みに対し、AIによる審査の場面では以下の特殊性を考慮する必要がある。

## ➤ ルールベースの審査を行っていないこと

- AIは評価対象者に関する要素を総合的に考慮し、個別にリスク要因の有無や程度を判断しているに過ぎず、「センシティブ属性を不利に評価する」といった評価基準に沿って審査しているわけではない。  
→AIの統計的バイアスの公平性を評価する指標（公平性評価指標）を踏まえつつ、従来の判断枠組みを適用していくことが考えられる（次頁【論点5-1】）。
- 前述のとおり、AIの活用により多様な要素を考慮した精度の高いリスク分析が可能となることが期待される。センシティブ属性が考慮要素に含まれたり、低評価を受けやすい傾向にある場合、裁判例の基準（前記8頁）に照らし、その合理的な根拠を説明できることや、センシティブ属性を過度にリスク要因と評価しないこと等が重要と考えられる（後記13頁）。

## ➤ 審査の過程が重層的であること

- 従来は単一の事業者内における取引条件の設定や当該条件に基づく取引拒絶の是非が争点となつたため、専ら故意の不法行為の成否が論じられてきた。
- これに対し、AIを用いた審査では、個々の出力においてどのような要素が重視されているのか（センシティブ属性が大きな影響を与えてるのか、そうではないのか）を外部から把握することが困難な場合が少なくない。  
→不公平な取引拒絶について自覺的でない場合（過失が論点となる場合）が想定され、AI開発者・提供者の設計上・説明上の注意義務と、AI利用者の利用上の注意義務の双方が論点となる。

## 【論点5-1】Aの責任（1/3）—設計上の措置①

- AI開発者・提供者であるAについては、データやアルゴリズムのバイアスが差別的な取引拒絶を生じないようにするため設計上どのような措置を講じるべきかが論点となる。従来の裁判例の枠組み（前記8頁）に照らし、以下(1)～(4)の観点から検討されると考えられる。

▶ まず、AIの統計的バイアスの公平性を評価する指標として下記のようなものがあり、下記(1)～(4)の検討の前提となる。

| 指標                                          | 視点                                                                        | 具体例                                                            |
|---------------------------------------------|---------------------------------------------------------------------------|----------------------------------------------------------------|
| Equal Opportunity<br>(機会の公平性)               | 実際に合格すべき（正解ラベルが1の）人をAIが合格と判定できた割合（True Positive Rate, “TPR”）が属性によらず同じであるか | 男性のTPR=80%、女性のTPR=60%であれば、女性は本来合格すべき人の20%分が不合格とされている           |
| Calibration差<br>(確率予測の信頼性)                  | AIが付与した予測スコアが、属性ごとに実際の合格確率と一致しているか                                        | 男性のスコアを0.8と予測した場合、実際に80%の男性が合格するが、女性のスコアを0.8と予測した場合、60%しか合格しない |
| Demographic Parity <sup>5</sup><br>(合格比率の差) | AIが合格と判定する割合が属性ごとに同じか                                                     | AIが合格と判断する割合が男性60%・女性40%であり20%の乖離がある                           |

### (1) 当該取引における相手方の属性の重要度

- 例えば想定事例5のように一定の長期間に亘って継続する賃貸借契約を締結する場面では、賃借人の属性は一般に重要な考慮要素であり、取引相手方の選択に係る賃貸人の裁量は尊重される。
- 他方、賃貸不動産は広く一般に提供されており、閉鎖性の高い市場とは言えず、目的の合理性や手段の適切性は問われ得る。

### (2) 不利益の程度

- 賃貸物件の契約は生活の基盤となる重要な機会であり、却下された場合の不利益は一般に大きいと評価され得る。
- 同一のAIサービスが用いられる範囲によっても不利益の程度は異なる。
  - 例えば、複数の不動産事業者が用いる汎用的なサービスか、或いは個別のオーナーのために開発されたサービスか

<sup>5</sup> 応募者集団に収入水準等の実力差がある場合、DPの均等を強制すると逆差別や業務上の弊害を生ずる可能性があるため、指標としての位置づけには留意が必要である。

## 【論点5-1】Aの責任（2/3）—設計上の措置②

- 仮に公平性指標において一定の乖離が認められたとしても、そのような設計に目的の合理性や手段の適切性が認められれば、権利侵害とは評価されない。

### （3）目的の合理性

- ✓ AI利用者の事業規模、取り扱う不動産の性質や重要性等に鑑み、センシティブ属性も踏まえた評価を行う必要性が高い場合、目的の合理性を基礎づける方向に傾く。外国人の住宅ローン審査において、返済可能性を確保する目的との関係で、定型的な条件を付す必要性の高さ等を考慮した前記9頁・参考裁判例②が参考となる。
- ✓ 開発上リスクを低減させる観点では、学習させる属性を慎重に決定し、極力センシティブ属性の学習を避けることや、代替する考慮要素を検討すること（AI事業者ガイドライン別添87頁参照）は、AIモデルが直接的にセンシティブ属性を不利に評価することを避けつつ、学習に用いた属性の根拠を説明可能にしておく点で、目的の合理性を裏付ける方向に評価されると考えられる。

例：公平性指標のうちEqual Opportunityにおいて、女性のスコアが男性より低い場合であっても、それが性別を直接理由とするものでない代理変数（例えば総合職やパートタイムなどの就労形態）に由来する場合であって、当該代理変数が家賃滞納リスクとの関係で顕著な関連性を有する場合には、目的の合理性が認められる方向に傾くと考えられる。

### （4）手段の適切性・代替手段の有無

- ✓ 手段の適切性は、センシティブ属性を不利に評価する傾向による対象者への不利益の程度も踏まえつつ、講じた緩和措置の内容、より不利益の小さい他の代替手段の有無等によって検討される。
- ✓ このような判断枠組みに照らし、必要に応じてAI事業者ガイドラインにおいて求められる以下のような措置（後記16頁も参照）を講じていた場合は、手段の適切性を裏付ける方向に評価されると考えられる。
  - データの比率やラベルの見直しによる緩和措置を講じたこと（それでもなお一定のバイアスが残存したこと）
  - センシティブ属性へのバイアスがより小さくなるよう他の考慮要素を採用したこと
  - 公平性に関するペナルティ項を追加した正則化の実施
  - バイアスによって合否に影響が生じやすい範囲を中心に入による判断（HITL）が必要な範囲を合理的に設計すること 等

## 【論点5-1】 Aの責任（3/3）—AI利用者が適切な利用を行うための措置

- データやアルゴリズムに一定の統計的バイアスが含まれていたとしても、利用段階で人が独立かつ公正な判断を下すことで、AIのバイアスが不当な取引拒絶に繋がることを回避できる。
  - AI開発者・提供者の注意義務違反の可能性を低減する事情として、AI利用者による適正利用に配慮する上での以下のような要素も考慮され得る。
- AIの利用段階で人が最終的な判断を行うことは、AIのバイアスが不公正な取引拒絶に繋がることを防止するために有効であり、AIの利用過程において一定の範囲で人による判断が行われていることは、手段の適切性（前記13頁(4)）を裏付け、注意義務違反の可能性を低減する事情として考慮される。
- その際、担当者が事実上AIの判断を機械的に援用するような運用を行うのではなく、公正に審査の判断を行っていたことが重要と考えられる。AI開発者・提供者は、AI事業者ガイドラインにおいて求められる以下のような措置（後記16頁も参照）を講ずることで、更に手段の適切性が認められる可能性が高まると考えられる。
    - ・バイアスのリスクの程度や性質を明確に説明すること
    - ・バイアスが合否判断に影響しやすい範囲等を踏まえ、人による判断を行うべき範囲を慎重に設計すること
    - ・可能な範囲でAIが判断の重要な根拠とした属性を表示したり、人が判断を行った後にAIの分析結果を表示するなど、UI上の工夫を行うこと 等

## 【論点5-2】Bの責任

- AI利用者は、審査の主体として差別的な取引拒絶を生じないようにする注意義務を負い、その判断に当たっては主に以下のような要素が考慮されると考えられる。

### ➤ データやアルゴリズムのバイアスに関する調査・確認

- 取引審査を公正に行なうことは、AI利用者が第一次的に（潜在的）顧客に対して負う義務であり、AI利用者は、採用しようとするAIサービスのバイアスの内容や程度について確認することが重要と考えられる。
- 前記12-14頁の考慮要素に配慮したAIサービスをAI利用者が用いる場合、そもそも取引拒絶が権利侵害には当たらないか、過失が否定される可能性が高まる。また、AI開発者・提供者がリスクに関し事実と異なる説明を行った場合など、AI利用者がやむを得ずバイアスのリスクを認識することができずにAIを導入したような場合には、過失が否定される可能性も想定される。

### ➤ バイアスのリスクに応じた人の判断の介在

- 前頁のとおり、AI利用者の利用過程では、人が最終的な判断を行うことで不公正な取引拒絶に繋がることを防止し得るが、AIの出力を機械的に採用するのではなく、人が公正な判断を行っていることが重要と考えられる。
- AI利用者は、AI事業者ガイドラインにおいて求められる以下のようないくつかの措置（次頁）を講じることにより、人が判断を行う上で適切な注意を払っていたことを示すことが可能と考えられる。
  - ・バイアスのリスクや人が判断を行うべき範囲等に関し、AI開発者・提供者から説明を受けたこと
  - ・人が判断を行う場合の客観的な評価基準を設定すること
  - ・人がAIの判断を覆した履歴やその理由を記録していること 等

※想定事例5とは異なる論点として、AI利用者がバイアスを惹起するようなインプットを入力した場合（例えば担当者が自由記述欄に特定の属性を揶揄するような主觀的記述を加えた場合）、これによって対象者のスコアが低減するバイアスが生じる可能性もあり、このことが別途不法行為を構成する余地もあり得る。

## (参考) 公公平性に関するAI事業者ガイドライン上求められる措置

- 公平性に関する共通の指針：①AIモデルの各構成技術に含まれるバイアスへの配慮、②必要に応じた人間の判断の介在（本編p16-17）

|                                                                                                   | 求められる措置                                                                                                                                                                                                                                                                  | 該当箇所                                                                             |
|---------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| 不動産業者B<br>       | <ul style="list-style-type: none"> <li>プロンプトに含まれるバイアスへの配慮</li> <li>必要に応じてAIモデルのバイアスに関する情報を収集</li> <li>バイアスに配慮しつつ<b>人間の判断を介在させる</b></li> <li>出力結果を合理的な範囲で関係者に提供</li> </ul>                                                                                                | 本編p38-39<br>本編p38-39・別添p155<br>別添p155<br>本編p38-39                                |
| AIサービス開発事業者A<br> | <ul style="list-style-type: none"> <li>データの公平性や参照する情報、連携する外部サービス等のバイアスを検討</li> <li>AIモデルの入出力及び判断根拠を定期的に評価</li> <li>バイアスの発生をモニタリングした上で必要な措置を検討</li> <li><b>UIやビジネスプロセスの設計</b>によるAI利用者又は業務外利用者へのバイアスの影響を検討</li> </ul>                                                      | 本編p35-36・別添p128-130<br><br>本編p35-36<br>本編p35-36・別添p128-130<br>本編p35-36           |
| AI開発者                                                                                             | <ul style="list-style-type: none"> <li>憲法14条1項等を考慮した<b>センシティブ属性の決定</b></li> <li><b>データに含まれるバイアス</b>に対する学習前・学習中・学習後を通じた緩和措置（代表性確保、潜在的バイアスの検知と除去等）</li> <li>公平性指標（メトリクス）によるアルゴリズムの<b>公平性評価</b></li> <li><b>人間による判断</b>を行うべき範囲の設計</li> <li>モデルの特性等に関する関係者への透明性確保</li> </ul> | 別添p87<br>本編p31・別添p87-88<br><br>本編p31・別添p94<br>本編p31・別添p95<br>本編p31-32・別添p103-104 |

※本表では重要な要素を概括的に抽出しているため、詳細については該当箇所を参照されたい。

## 【論点5-3】因果関係（1/2）

- AIのバイアスは統計的なものであるところ、因果関係の評価が困難な場合があり得る。
- 統計的なリスクがあることだけでは、それが取引拒絶の直接の原因になったとは断定できず、事実的因果関係は認められない。  
e.g.) 仮にAがバイアス抑制策を講じ、例えばEqual Opportunityの男女差が10ptから5ptに軽減したとしても、改善後のアルゴリズムを用いた場合に、Vが審査に合格していたとは限らない。
- 他方、統計的にリスクが高いことは因果関係を基礎づける一つの事情とはなり得る。参考裁判例⑦のように、他の間接事実と組み合わせることにより因果関係が肯定される場合もある<sup>6</sup>。

### 参考裁判例⑦：X線照射事件（最判昭和44年2月6日民集23巻2号195頁）

- レントゲン照射と皮膚癌との間の因果関係について、照射と癌（特に皮膚癌）の発生との間に統計上の相関があること、当該事件における照射の時期・量・回数および部位、被害者の皮膚癌が照射部分についてのみ発生したこと等の諸事実を踏まえ、因果関係を認めた。
- 参考裁判例⑦では、統計的な相関性の高さに加え、(i)行為と結果との間の時間的近接性、(ii)他原因の不存在、(iii)原因と結果の間の関連性が科学的・客観的に説明しやすく、不自然な因果経過が存在しないこと等が重視されている。
- 例えば想定事例5においては、Vの所得や勤務先の信用度等の経済条件が十分に高く、センシティブ属性が影響しなければ信用スコアが高かったはずであること等が示されれば、因果関係も認められると考えられる。

<sup>6</sup> 因果関係の立証は、一点の疑義も許されない自然科学的証明ではなく、経験則に照らした「高度の蓋然性」の証明で足りるとした最判昭和50年10月24日民集29巻9号1417頁（レンバール事件）も参照。

## 【論点5-3】因果関係（2/2）

- 因果関係が存在する「高度の蓋然性」の認定が困難な場合、以下のような枠組みを用いるべき場合があり得るか。
- 確率に応じた損害額の認定：原告の被った損害に対し、損害が不法行為に起因する確率を乗することにより損害額を定める考え方

参考裁判例⑧：水俣病訴訟(1)（東京地判平成4年2月7日判時平成4年4月25日号3頁）

- 水俣病の認定申請を棄却された者で、死亡後に病理学的に水俣病と診断された例は71例中19例（27%）と少なからず存在するところ、因果関係の証明につき高度の蓋然性を要求すると、医学の限界による負担を一方的に原告に課すことになり、損害の公平な分担に反することから、水俣病であることの「相当程度の可能性」が認められる場合、その可能性の程度を損害賠償額（慰謝料）の算定に反映させるのがより妥当と考えられると判示した。

参考裁判例⑨：水俣病訴訟(2)（大阪地判平成6年7月11日判時1506号5頁）

- 行政上の水俣病の認定基準に合致する原告らについては「高度の蓋然性」を認めつつ、認定基準に合致しない原告らについて、水俣病であることの可能性を症状の組合せごとに「40%・30%・20%・15%」の4段階に区分し、基準額にその割合を乗じるという按分方式を採用した。

課題：上記各判決は、公害事案であることや医学的な限界が前提となっており、どこまで射程が及ぶか。また、損害が不法行為に起因する可能性をどのように合理的に推計するか。

- 代替法益の侵害：最終的な被侵害法益でなく、その継続可能性や期待権が侵害されたと捉える考え方

参考裁判例⑩：相当程度の生存可能性（最判平成12年9月22日民集54巻7号2574頁）

- 因果関係の存在が証明されなくとも、「医療水準にかなった医療が行われていたならば、患者がその死亡の時点においてなお生存していた相当程度の可能性の存在が証明されるときは、医師は、患者に対し、不法行為による損害を賠償する責任を負う」と判示した。

参考裁判例⑪：適切な治療を受けることへの期待権（大阪地判平成23年7月25日判タ1354号192頁）

- 赤十字血液センターへの連絡過誤により輸血手配が約30分遅延した事案において、死亡との間の因果関係は否定したものの、適切な医療を受ける期待権の侵害を理由として慰謝料の損害賠償請求を認容した。

課題：合理的な代替法益をどのように設定するか。

## 【論点5-4】立証上の論点

- 想定事例5との関係では、Vが各要件についての主張立証責任を負うが、文書提出命令等を通じ、AやBからもバイアスの内容等に関する説明が求められる可能性がある。
- AやBが適切な注意を払っていたことを示す観点では、以下のような証拠が重要と考えられる。

- AIのバイアスは外部から検証することが困難であり、Aが保有するものとしては以下のようないる資料が重要と考えられる。
  - ・ 公平性評価レポートその他AIモデルのバイアスの概要に関する書面
  - ・ バイアスを低減するために講じた措置の概要に関する書面
  - ・ 人が審査に関与すべき範囲等、運用上の留意点を記載したAI利用者向けの説明文書 等
- また、Bが保有するものとしては以下のようないる資料が重要と考えられる。
  - ・ AIを導入する際にバイアスのリスクに関する検討内容を記した書面
  - ・ 人が判断を行う場面での留意事項その他利用上の体制・ガバナンスに関する書面 等

1. 第2回研究会でのご意見
2. 想定事例5：取引審査AI
- 3. 想定事例6：自律走行ロボット（AMR） - 前半**
4. 想定事例6：自律走行ロボット（AMR） - 後半

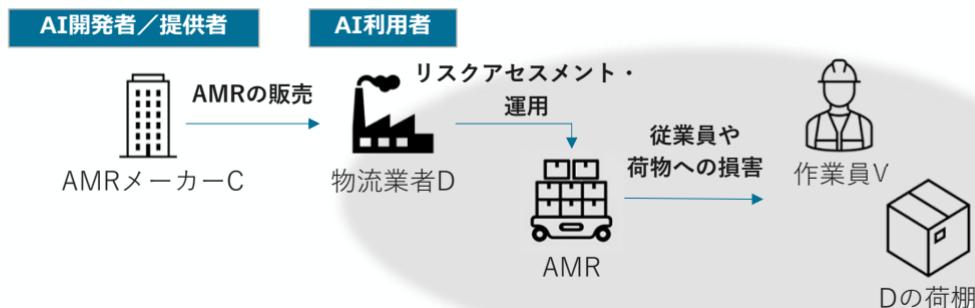
## 想定事例6：自律走行ロボット（AMR）

- AMRメーカーCは、倉庫や工場等の現場で稼働する台車型の自律走行ロボット（Autonomous Mobile Robot, “AMR”）を製造・販売している。当該AMRは重量300kg・可搬重量500kg、最高速度1.2m/秒（人が歩く程度の速度）で稼働し、従来の台車やフォークリフトに代わって工場内の荷物を迅速に運搬することができる。
- 物流業者Dは、CのAMRの性能や自社倉庫の環境を踏まえてリスクアセスメントを実施した上、人との協働を行う現場へのAMRの導入を決定し、購入した。当該AMRは長期間に亘り安全に稼働したが、ある日、多量の荷物を運搬中に以下の事故が発生した。

【分岐a】AMRの前方にいた従業員Vに衝突し、負傷させた（後記28頁以下）。

【分岐b】AMRの機体に、ログやアラートを分析しトラブルシューティングを行うAI（言語モデル）とその出力を表示するタッチパネルが搭載されていたが、バッテリの温度上昇に関する誤った診断を作業員に伝えた結果、機体が発火し延焼した（後記34頁）。

【分岐c】AMRの引渡し後、運用過程におけるアップデートにおいてソフトウェアに重大なバグが混入し事故が生じた（後記35頁）。



# 論点整理

- 下記のとおり、AMRのメーカーであるCの製造物責任及び運用者であるDの過失が論点となり、それぞれがどのような事項について責任を負うかが重要な観点となる。

- 【分岐a】人身損害の場合
  - ・ 本想定事例のように従業員が負傷した場合、治療費等の一定の損害は労働者災害補償保険（労災保険）によって填補される。これに対し、慰謝料や後遺症逸利益等が生じた場合には労災保険の補償対象とならず、雇用主であるDの責任（安全配慮義務違反ないし過失責任）やロボットのメーカーであるCの責任（主に製造物責任）が論点となる。
  - ・ 従来、例えばフォークリフトによる事故の場面では、運転手の過失責任を認定し、雇用主に対して使用者責任を追求するという解決が取られていた（後記37頁・参考裁判例⑫）。これに対し、自律走行ロボットを運用する場面では、ソフトウェアが自律的に行う運転判断について過失責任を問うことができないため、CやDがそれぞれどのような範囲で責任を負うかが問われる。
- 【分岐b】財産的損害の場合
  - ・ Dの財産である荷物が焼損していることから、ロボットのメーカーであるCの責任（主に製造物責任）が論点となる。
  - ・ 財産的損害についても製造物責任は適用されるから、責任が生ずる要件は【分岐a】の場合と共通する。Dに運用上の過失があった場合には、過失相殺に基づく損害賠償額の減額として考慮される。
- AIを含むソフトウェアによって稼働する様々な機械やロボットの登場が予測される中、これらの機械の挙動によって事故が生じた際の責任関係については、裁判例の蓄積がなく予測可能性に乏しい。
- ロボットの挙動により損害が生じた場合、本想定事例のように製造業者（メーカー）と運用者（ユーザ）の責任が争点となるケースは多いと考えられ、それぞれがどのような責任を負うかという観点も重要となる。

7 自動運転車については特別不法行為法の適用を含め議論が進んでいるが（自動運転SWG報告書、自賠責保険検討会）、今後はその他のロボットの更なる普及も見込まれる。

8 なお、次頁のとおり、事業所等においてロボットを運用する者（本想定事例のD）は、労働安全衛生法に基づくリスクアセスメントの義務（同法28条の2）等を負う。JIS規格においてもリスクアセスメントの実施が求められている。

# AMRの安全性に関する技術及びリスクアセスメント

## ➤ 安全系：人や物との衝突を防止するための機能

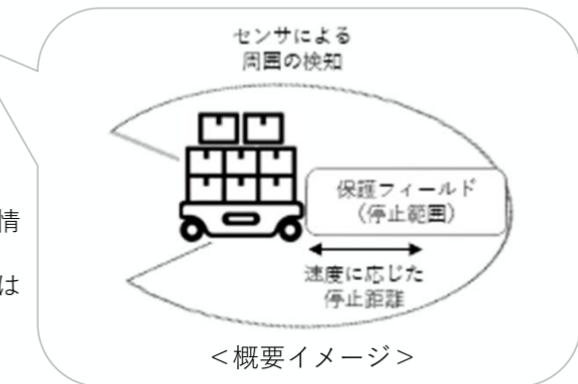
- 必要な機能は個々の製品や稼働現場において様々であるが、標準・規格としては、ISO 3641-4やそのJIS整合規格であるJIS D 6802等において、以下のような要求事項及びPLr（要求パフォーマンスレベル）が定められている。
  - レーザースキャナ等の検知センサ
  - 速度制御・監視機能
  - 速度や緊急停止距離に応じた保護フィールド（停止範囲）の設定
  - その他の安全機能（ブレーキ系、非常停止装置、警報等）
  - 保守的な走行条件を前提とした人検知試験

## ➤ 運用系：円滑な走行のための機能

- SLAM (Simultaneous Localization and Mapping) …センサで取得した環境情報から壁や障害物などの地図を作成しつつ、AMR自身の位置や向きを推定
- 3Dカメラ及びAI画像認識…人や物の区別が可能となり、人を認識した場合には予め減速する等、より柔軟な動作判断が可能となる

## ➤ ユーザによるリスクアセスメント

- 工場や倉庫においてAMRを運用するユーザはリスクアセスメントを行うべき立場にあり（労働安全衛生法28条の2）、具体的にはJIS B 9700 (ISO 12100)や JIS B 9705-1(ISO 13849-1)等を参照しつつ、概ね以下の手順により実施される<sup>[1][3]</sup>。
  - 仕様の決定、危険源の特定
  - リスクの見積り・評価
  - リスク低減措置：本質的対策（走行速度を下げる等）、工学的対策（付加的なセンサの追加等）、管理的対策（走行ルール等の策定や作業員への教育等）、個人用保護具の使用 等
  - 実施状況の記録・見直し



参考：

[1] 中央労働災害防止協会「技術革新に対応した機械設備の安全対策の推進事業 報告書」令和3年10月31日 (<https://www.mhlw.go.jp/content/000830713.pdf>)、令和7年11月25日最終閲覧)

[2] 工場・物流DX導入センター「AMR制御システム完全ガイド～どう動く？何が違う？現場で役立つ判断ポイント」令和7年6月29日 (<https://factory-dx-center.com/amr-control-system-guide/>)、令和7年11月25日最終閲覧)

[3] 株式会社リンクス「AMR導入の失敗を防ぐ！Part3「安全対策」を考える～法令解説から現場の運用まで～」令和7年8月20日 ([https://linx.jp/product/irayple\\_amr/article/009/](https://linx.jp/product/irayple_amr/article/009/))、令和7年11月25日最終閲覧)

# 製造物責任法における「欠陥」の類型

- 製造業者等は、製造・加工・輸入等をした製造物に欠陥があり、他人の生命、身体又は財産を侵害したときには、これによって生じた損害の賠償責任を負う。一般に、欠陥は以下の三類型に整理して説明される。
- AMR等のロボットについては、機体の不具合や設計不備（例えば発火の危険性等）が従来と同様に問題となるほか、ソフトウェアの挙動によって生じる事故についてどのように欠陥判断を行うべきかが論点となる。



製造物が設計・仕様どおりに作られず安全性を欠く場合

設計上の欠陥

製造物の設計段階で十分に安全性に配慮しなかったために、製造物が安全性に欠ける結果となった場合

指示・警告上の欠陥

有用性ないし効用との関係で除去し得ない危険性が存在する製造物について、その危険性の発現による事故を消費者側で防止・回避するのに適切な情報を製造者が与えない場合

安全措置について設計どおり作られた製造物とそれに欠ける当該製造物との間の相違という事実認定が重要となる

多様な設計上の選択肢の中から製造者が選択した結果の適否に関する規範的な評価が重要となる

## 【論点6-1】欠陥該当性に関する判断基準

- 欠陥の判断において重視されてきた要素を踏まえると、AMRは以下のように評価される。
- 欠陥の判断においては、(1)製造物の特性、(2)通常予想される使用形態、(3)引渡し時期その他の事情を考慮することとされ（製造物責任法2条2項）、より具体的には以下のような事情が考慮される<sup>9</sup>。

|            | 指標               | 視点                                                                                                                                                                                                               |
|------------|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (1)に関連する要素 | 製造物の表示           | <ul style="list-style-type: none"><li>・衝突の危険性等は外部から認知しやすいが、安全性に影響する技術的な設計内容や予想外の挙動をしやすい状況等については外部から推知することが困難であり、リスクに関する慎重な説明が求められやすい</li></ul>                                                                   |
| (1)に関連する要素 | 製造物の効用・有用性       | <ul style="list-style-type: none"><li>・機動的かつ迅速に運行することで従来の台車等での作業に比べ大量の荷物の搬送等が可能となる</li><li>・センサ等を過度に冗長に設計し停止頻度が多くなると、円滑な運行という本来の目的自体が阻害される</li></ul>                                                             |
| (1)に関連する要素 | 価格対効果            | <ul style="list-style-type: none"><li>・異種／複数センサを併用する等、冗長設計を多重にすることで安全性を向上させるカスタマイズは可能であるが、求められる安全設計は利用環境によって様々であり、ユーザのリスクアセスメントを踏まえて仕様が決定され、当該仕様を踏まえてコストが定まる</li></ul>                                           |
| (2)に関連する要素 | 使用者による損害発生防止の可能性 | <ul style="list-style-type: none"><li>・利用環境は個々の工場や倉庫の状況に応じて様々であり、ユーザの管理下にあるため、ユーザによるリスクアセスメントを踏まえてロボットの速度制御等がカスタマイズされており（前記23頁）、レイアウトの調整や変更も行われる</li><li>・利用環境の設計や作業員への教育指導により、積極的に事故回避措置を行うことが予定されている</li></ul> |
| (2)に関連する要素 | 合理的に予想される使用形態    | <ul style="list-style-type: none"><li>・一定の危険行動や想定利用環境から逸脱した稼働は予想されるが、自律走行車両に関する教育を受けた者との協働が想定されており、作業員による合理的な対応も期待しやすい</li></ul>                                                                                 |
| (3)に関連する要素 | 引渡し時における技術水準     | <ul style="list-style-type: none"><li>・現時点では前記23頁のような技術が用いられることが一般的</li></ul> <p>※なお、実際にAMRが稼働する状況等を踏まえ、ソフトウェアアップデートが頻繁に行われるのが通常であり、この点を踏まえた欠陥判断について後記35頁参照。</p>                                                 |

9 消費者庁「製造物責任(PL)法の逐条解説 第2条（定義）」平成30年9月

([https://www.caa.go.jp/policies/policy/consumer\\_safety/other/product\\_liability\\_act\\_annotations/pdf/annotations\\_180907\\_0003.pdf](https://www.caa.go.jp/policies/policy/consumer_safety/other/product_liability_act_annotations/pdf/annotations_180907_0003.pdf))、令和7年11月25日最終閲覧)

考慮要素は他にも存在するが、AMRとの関係で特に重要と考えられるものを抽出している。

## 【論点6-1】欠陥該当性に関する判断基準

- 欠陥該当性に関する判断基準については以下のような考え方がある。
- 自動運転車等の自律的に稼働する製造物について、どのように安全性を判断するかが議論されている。

|                          | 概要・根拠                                                                                                                                                                                   | 課題                                                                                                                                                                                                 |
|--------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (1) 総体的・統計的な安全性を基準とする考え方 | <ul style="list-style-type: none"><li>人の判断や行動と比較し、ロボットの自律的な判断の方が確率的・統計的に安全か否かを基準とする考え方。</li><li>事故発生の蓋然性は欠陥判断の主要な考慮要素の一つと考えられている<sup>10</sup>。</li></ul>                                | <ul style="list-style-type: none"><li>消費者としては個別状況における危険性や事故回避可能性も重視すると考えられ、統計的な安全性を絶対的な基準とはし難いという指摘もある。</li><li>統計的な安全性をどのように比較するかが難しい場面も想定され、AMRのように普及初期にあるシステムではデータを取ることが困難なケースも少なくない。</li></ul> |
| (2) 個別状況における安全性を基準とする考え方 | <ul style="list-style-type: none"><li>事故が生じた個別の状況を踏まえ、発揮すべき安全性を備えていたかどうかを評価する考え方。</li><li>自動運転車の文脈では、一般的な人の運転技能と同様の安全性を求める考え方などが主張されている。</li><li>消費者の期待保護の観点からは個別状況を評価すべきである。</li></ul> | <ul style="list-style-type: none"><li>AMRのように人が従来運転操作をしていなかった機械の場合、どのように個別状況の安全性を評価するかが論点となる。</li></ul>                                                                                            |

- (1)及び(2)については、技術の進展に伴い、人の能力のみならず別のシステムの安全性が比較対象となることも考えられる。
- また、(1)及び(2)における安全性の検討に当たっては、規制法や標準・規格等の安全基準に適合していることも有力な考慮要素となる<sup>11</sup>。
- 各基準は、製造物の性質や損害発生の状況等に応じて適切な場面が異なることも想定される。  
→基準の明確性や消費者の期待水準を考慮しつつ、後記【論点6-2】や【論点6-4】のとおり、場面に応じた判断枠組みの検討が必要

10 前記25頁注9・消費者庁9頁

11 前記25頁注9・消費者庁20頁、後記37頁・参考裁判例⑬

## 【論点6-2】分岐a—検討の方向性

▶ 前記25頁の考慮要素（特にAMRの特徴である個別の利用環境に即したカスタマイズ性やユーザ側での損害防止措置の可能性）や、前記26頁の判断枠組みを踏まえると、以下のような検討の方向性が考えられる。

### ① 製造上の欠陥

- ・ 合意された仕様や設計が重大なバグ<sup>12</sup>などの不具合により発揮されなかつたような場合、製造上の欠陥が論点となる。

### ② 設計上の欠陥

- ・ 個別状況の安全性は、AMRの設計（重量・速度・停止距離等の安全に関わる設計）、その他合意された設計やカスタマイズの内容を踏まえ、当該事故を回避可能な代替設計が存在したか否かを評価する。

a. 想定利用環境における通常の利用過程でのリスクについては、技術的に可能な限りの安全設計が求められやすい。

b. 合理的に予見される誤使用や想定利用環境の逸脱についても、一定のフェールセーフ等の設計上の配慮が求められる。

- ・ 他方、稼働現場はユーザの管理下にあり、利用過程や利用環境側での安全確保措置も重要となる。そのため、ユーザのリスクアセスメントにおいて合理的に予見可能な事象であり、ユーザ・作業員による事故回避の実効性・容易性が認められる場合や、製造者から禁止事項と指定された内容に反した使用方法を取った場合は、欠陥が否定される方向に傾く。

- ・ 標準・規格としてはISO及びJISにおいて詳細な安全基準が定められており、一応の基準となるが、最終的には個別の場面における必要な安全性を満たしていたか否かで評価される。

### ③ 指示・警告上の欠陥

- ・ AMRはソフトウェア及びハードウェアの複雑な統合によって構成されるところ、ユーザが安全性能を適切に発揮し合理的な回避行動を取る前提として、性能限界や残留リスクの明確な説明が求められる。

12 ソフトウェアが複雑になればなるほどバグが存在しない状態を実現することは非現実的であり、全てのバグが「通常有すべき安全性」を欠くと評価されるものではない。また、AIが意図通り挙動しなかったことを以てソフトウェアのバグとは評価し難い点について、後記29頁・注13参照。

## 【論点6-2】分岐a—事例検討（1/2）

- 想定事例6では、衝突の態様に様々なパターンが考えられ、それぞれ評価が分かれる。
- 具体的な評価は個別の事故の状況によっても異なるが、大まかな傾向としては以下のように整理されるのではないか。

**【分岐a-1】開発過程に起因する事故：**VはAMRの右側方から歩行して走行経路に侵入したところ、合意された設計上は衝突前に停止が可能であったが、ソフトウェアに特定の条件下で停止距離を過小に計算してしまうバグがあったため、停止が遅れ衝突した。

- ✓ 上記事例では合意された設計としても事故を回避可能であることが想定されていたところ、ソフトウェアのバグにより事故が発生していることから、製造上の欠陥ないし設計上の欠陥に当たる可能性が高い（前記27頁参照）。

**【分岐a-2】利用環境や利用方法に起因する事故：**AMRは表面の汚損によりセンサ感度が低下するおそれがあることから、CはDに対し、当該リスクと共に定期的なメンテナンスの必要性やその方法を説明していた。しかし、Dは使用過程でメンテナンスを怠っており、機体に汚れが付着していた。その結果、AMRの走行経路上で機体の存在に気づかず作業していた作業員の検知、警報及び制動が遅れ、衝突した。

- ✓ 誤使用による事故を全て防止するための設計は求められないが、上記は合理的に予見される誤使用と考えられ、一定のフェールセーフ（センサ感度の異常を検知・警告する仕組み等）を備えていたかは問われる。
  - なお、複数のセンサやカメラを併用していた場合、事故が回避可能であった可能性も想定されるが、当初合意された利用環境や設計内容は尊重され、これを大きく逸脱するコストを前提とした代替設計は、合理的な代替設計とは認められづらい。
- ✓ 機体表面の汚れによるセンサ感度の低下リスクや必要なメンテナンス方法を説明している場合、Dにおいて上記事故の発生は合理的に予見することができ、かつメンテナンス等の対処も実施可能である。このようなユーザ側による事故回避の容易性や、そのための十分な情報が提供されていたことは、設計上の欠陥及び指示・警告上の欠陥が否定される方向に傾く。
  - Dが適切なメンテナンスを怠っていた場合、過失責任を負う可能性がある。

## 【論点6-2】分岐a—事例検討（2/2）

### 【分岐a-3】設計上の限界と利用方法が競合して生じる事故：

- CのAMRは、JISに準拠した基本的な衝突防止機能に加え、3Dカメラ及び画像認識AIを搭載し、側方から接近する作業員を検知して早期に警報を発しながら減速するように設計されていた。しかし、AIには検知限界があり、特異なシルエットについて検知精度が下がること等のリスクはCからDに説明されていた。
- Vは、作業場の臨時トラブルの対応のため巨大な脚立を肩に担いで横向きに持ちながら移動していたところ、AMRに搭載されたAIがVのシルエットを人であると検知できず、警報や減速を行わなかった。Vは更にAMRに接近し、その前方を横切ろうとした際、AMRが衝突した。
  - ✓ 事故の原因は、(1)基本的な衝突防止機能の動作と、(2)画像認識AIの検知失敗に分けられる。
  - ✓ (1)については以下のとおりと考えられる。
    - このような安全設計がソフトウェアのバグ等により適切に機能しなかった場合、製造上の欠陥ないし設計上の欠陥が疑われる。
    - 安全設計が機能していたものの、Vがあまりにも突然接近したことで事故が生じたような場合、製造上・設計上の欠陥を否定する方向に傾く。
  - ✓ (2)については以下のとおりと考えられる。
    - 設計の適切性の観点では、以下の点が重要と考えられる。
      - 想定事例3で議論を行ったAI開発者によるAIガバナンスの実施状況（開発上合理的に可能な範囲でAIの精度を向上させること、実運用に近い環境下で精度や安定性を検証すること等。第2回研究会事務局資料13頁）が問われる<sup>13</sup>。但し、(1)のようなフェールセーフが備わっていることや、例外的なシチュエーションについてAIの精度を高めることの困難性は考慮される。
      - その上で、Dにおいて誤検知が生じやすい荷物の運搬を避けるなどの事故回避措置を講じることが可能であったと評価されれば、設計上の欠陥が否定される方向に傾く。
      - 指示・警告の観点では、AIの検知が成功しづらい状況や誤検知の場合のリスク等、Dがリスク軽減策を講じるために重要な情報を説明していなかった場合には、指示・警告上の欠陥に該当する可能性はある。
  - ✓ Cによる設計上の限界等についての説明を踏まえ、Dが事故回避のための合理的措置を講じていなかったと評価される場合、過失責任を負う可能性がある。また、V自身の行動が不用意な接近と評価される場合、過失相殺として損害賠償が減額される可能性がある。

13 なお、画像認識AIによって常に望ましい検知を保証することは性質上困難であるため、意図通り検知しなかったことを以てソフトウェアのバグとは評価し難いと考えられる。

## 【論点6-2】分岐a—事故を防止する上のD及びVの責任

- DはVに対し、安全配慮義務に基づく責任を負う場合がある。
- また、Vにも過失ないし何らかの不注意があれば、過失相殺として損害賠償額に反映される。

### ➤ Dの安全配慮義務

- ✓ DはVに対し、労働契約上の安全配慮義務を負う立場にあり、その一環として、AMRの導入に当たってリスクアセスメントを実施することが法令上求められている。
  - Dは労働者の安全と健康を確保する観点からリスクアセスメントを行うべき立場にあり、事業所における潜在的な危険源を特定し、そのリスクを評価し、適切な低減措置を講じるべきこととされている（労働安全衛生法28条の2）。
- ✓ 導入時や運用過程におけるリスクアセスメントを通じ、リスクを低減するために必要と判断された措置や体制構築を含め、事故回避のための合理的対策を講じることが安全配慮義務の内容を構成すると考えられる。

### ➤ Vの過失ないし不注意

- ✓ CやDが損害賠償責任を負う場合であっても、事故との関係でVに「過失」（民法722条2項）が認められる場合には、過失相殺として損害賠償額が減額され得る。
- ✓ ここでの過失は法的な注意義務（予見可能性を前提とした結果回避義務）に限らず、事故の発生や損害の拡大に寄与したと評価できるような不注意や軽率な行動も含まれると解されており、例えばDが策定したマニュアルや手順書からの逸脱等があった場合には過失相殺の対象となり得る。

## 【論点6-3】中長期的な視点—AIによる予測困難な挙動

- 今後、AIが動的に環境地図を更新したり経路を生成することにより、柔軟な環境への対応可能性が広がり得る。
- これに伴いAIが予測困難な挙動をした場合、欠陥判断をどのように行うべきか。

想定) CのAMRは、AIによる動的な地図更新や経路生成により現場環境の変化にも柔軟に即応できるように設計されていた。当該AMRは長期間に亘り安定して運用されていたが、突如、閉まっている扉へと直進し、衝突して扉を大破させた。事後の検証で、当該扉は通常は開放されていたところ、事故当時だけ閉められており、AIが地図上で「通り抜け可能」と判断したことが判明した。

- ✓ 検討の方向性としては前記27頁が妥当すると考えられるところ、上記の事故態様としては、ユーザにおいて予測や対応を行うことが難しく、危険性も高いため、個別の場面における安全性評価としては欠陥該当性が認められやすいと考えられる。
- ✓ 他方、AIに内在する不確実性のため、開発者において予見が著しく困難であったことが認められれば、欠陥を否定する事情として斟酌される。例えば以下ののような要素が重要と考えられる。
  - 設計上の観点として、AIが広範な制御権限を有することは適切であったか（環境への柔軟な対応という目的との関係で必須であったか）、誤判断に備えて適切なフェールセーフ<sup>14</sup>を備えていたか。
  - 想定事例3で議論を行ったAI開発者によるAIガバナンスの実施状況（開発上合理的に可能な範囲でAIの精度を向上させること、実運用に近い環境下で精度や安定性を検証すること等。第2回研究会事務局資料13頁）や、運用過程で生じた事故の再発を防止するためのアップデートを含め、事故を防止するための最善の手段を講じていたか。
  - 上記のような取組みにより、事故率が低い割合に抑えられていたか。
  - 生じた事故が、事前に把握していたリスクとは異質ないし特異なものであったか。
  - 運用過程での継続学習によってAIの性質が変化していたか。
- ✓ また、製造物の引渡し時点における科学又は技術に関する知見によっては欠陥の存在を認識することができない場合、損害賠償責任は生じない（製造物責任法4条1号・開発危険の抗弁）。従来の裁判例では開発危険の抗弁の適用を認めた事例はないが、不確実性が内在する科学技術であるAIに関しては、今後、開発危険と評価される事例が生じることはあり得る。

14 例えば、AIの判断の信頼度が低い場合には侵入を中止したり、侵入前に定点カメラ等で安全性を二重に確認する等の方法が考えられる。

## (備考) 他の種類のロボット

- 欠陥判断においては個々のロボットの特性や想定利用環境、運用者の性質等が重要な考慮要素となる
- AMR以外のロボットにおいては判断枠組みが変わりうる

### 工場等で稼働するAMR

- ・工場や倉庫等、閉鎖空間においてのみ稼働
- ・ユーザが法令上リスクアセスメントを行うべき立場にある
  - ・ユーザが稼働領域を全面的に管理することが可能
- ・ユーザの指導監督に服する作業員との協業
- ・人間による操作を前提としていない
  - ・詳細な規制や安全基準がない



### 他のロボットとの比較

- ・より開かれた空間で稼働する場合  
⇒子供や高齢者等との相互作用や想定外の環境変化等も考慮し、より一般的な安全性が要求される可能性
- ・ユーザによる適切な維持管理が必ずしも期待できない場合  
⇒設計上対応すべき範囲が広がる可能性
- ・ユーザに必ずしも運用上の知見がない場合  
⇒ユーザの知識に配慮した設計や、理解可能性を考慮した指示・警告が必要となる可能性
- ・従来人間が操作してきた機械と類似の形状  
⇒統計的な安全性や平均的な通常人の操作能力との比較により判断基準を明確化しうる
- ・規制や安全基準が存在する領域  
⇒基準の内容が欠陥判断の基準と近似する可能性

1. 第2回研究会でのご意見
2. 想定事例5：取引審査AI
3. 想定事例6：自律走行ロボット（AMR） - 前半
4. 想定事例6：自律走行ロボット（AMR） - 後半

## 【論点6-4】分岐b—AMRの自己診断のためのAI

- AMRに組み込まれたトラブルシューティングのためのAIについて「欠陥」が認められるか。

**【分岐b】** AMRの機体にはLLMを活用した自己診断AIが搭載され、ログやアラートを踏まえたトラブルシューティングを行い、タッチパネルを通じて診断結果や対処法を現場の作業員に伝える機能を有していた。あるときバッテリが高温となり、AMRが充電禁止の警報を発したが、上記AIは「温度センサの一時的な異常」と判断し、充電禁止のセーフガードを無効化する対応策を提示した。作業員は当該診断に従ったところ<sup>15</sup>、バッテリの過度な温度上昇が進行し、10分後に発火して倉庫内のDの荷物に延焼した<sup>16</sup>。

- ソフトウェアやUIの構成には複数の在り方が考えられるところ、どのような場合に自己診断機能が製造物の構成要素と評価されるか。
- ・ 消費者の期待に照らすと、例えば初期的な分析を行う軽量LLMや診断結果を示すUIやタッチパネルなど、自己診断機能に必要な要素の一部がAMRの機体に搭載されている場合には、自己診断機能全体がAMRの構成要素に当たると捉えるべきか。

| ソフトウェアの構成                                                                                                                                                                                    | UIの構成                                                                                                            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"><li>・ 機体に搭載したLLMで全ての推論を行う</li><li>・ 機体に搭載した軽量LLMで初期的な分析を行い、その内容を踏まえ、サーバ側のLLMで詳細な推論を行う</li><li>・ 機体はログやアラートの情報をサーバ・クラウドに送信するのみで、サーバ側のLLMで全ての推論を行う</li></ul> | <ul style="list-style-type: none"><li>・ AMRの機体のタッチパネル等で診断結果を表示する</li><li>・ AMRの機体とは独立のホストPC等で診断結果を表示する</li></ul> |

- 自己診断の誤りが製造物の欠陥を構成し得る場合、どのように欠陥の有無を判断するか。
- ・ 前提として、従業員が危険な設定変更をできないようにするためのアクセス制御等が講じられていれば上記のような事故は回避できる可能性が高く、このようなフェールセーフが講じられていたかが重要なポイントとなる。
  - ・ フェールセーフが機能せず、結果的にAIの出力の誤りが事故に至った場合、上記のようなトラブルシューティングは従来は技術者等の人間の判断・行動によって対処してきた場面であるところ、当該状況に直面した通常の注意力の技術者の判断と比較してAIの誤出力を評価することが考えられる（前記26頁(1)及び(2)参照）。
  - ・ その際、統計的・確率的な信頼性（前記26頁(1)）あるいは個別の場面における正確性（前記26頁(2)）のいずれが重視されるか。

15 仮に、Cとは別のメーカーEがAIを搭載したタッチパネルを製造し、CがこれをAMRに組み込んだ場合、当該タッチパネルについてのEの製造物責任や、製造物責任法4条2項に基づく抗弁が論点となり得る。

16 作業員がAIの出力に依拠して危険な操作を行ったことについては、Dの教育上・管理上の過失を踏まえた過失相殺や、作業員の過失をDとの関係で過失相殺として考慮し得るかも論点となる。被用者に業務遂行上の過失が認められる場合、被害者側の過失として過失相殺の対象になり得るが（最高裁判決昭和42年6月27日民集21巻6号1507頁等）、作業員がAIの誤った出力を信頼することがやむを得ないケースも想定し得る。自動運転車に関し、被害者が自動運転装置による制動を期待して行動した場合の議論について、佐野誠「多当事者間の責任の負担の在り方」藤田友敬編『自動運転と法』211頁以下（2018年、有斐閣）参照。

## 【論点6-5】アップデートに関する責任

- アップデートにおいてソフトウェアの重大なバグ等の不具合が混入した場合、欠陥の判断において考慮されるか。

**【分岐c】** CがAMRを引き渡した時点ではソフトウェアにバグが存在しなかったが、運用過程におけるアップデートにおいて、特定の条件下で停止距離を過小に計算してしまうバグが生じ、【分岐a-1】と同様、意図した設計であれば回避できた衝突事故が発生した。

- VはCに対し、少なくとも過失責任（民法709条）に基づく請求を行い得る。
  - VはCの過失（事故の予見可能性を前提に、Cがどのような内容の結果回避義務を負っていたか）を主張立証する必要があり、主張立証が可能な場合には問題なく損害の填補を受けることができる。
  - 他方、【分岐c】では論点を明確化するため事故原因を明示しているが、そもそも事故原因が明らかでない場合もある。製品に関する重要な資料はCやDに偏在しているところ、事故原因をVにおいて究明・特定し、過失の主張立証を行うことが困難な状況も想定される。
- Vは過失責任に加え、製造物責任に基づく請求も行うことができる。
  - 製造物責任に基づく場合、Vは過失でなく欠陥を主張立証することとなり、また欠陥推定法理（次頁）も活用し得る。
  - もっとも、欠陥判断の基準時は製造物の引渡し時<sup>17</sup>と解されているところ、引渡し後のアップデートにより生じた事象を欠陥判断において考慮し得るか。考慮される場合、どのような根拠が考えられるか。
    - ① 欠陥の判断は目的物の特性を踏まえて判断すべきであるところ、AMRは走行の効率性や安全性の向上のために頻繁なアップデートが予定されている。アップデートは引渡し時に予定されていたものとして、欠陥判断に織り込むべきか。
    - ② 製造業者自身が管理行為としてアップデートを行っているため、欠陥判断に織り込むことは不合理とは言えないのではないか。
    - ③ 上記のように引渡し時点でアップデートが前提とされる場合であれば、【分岐a-1】のみならず【分岐c】においても、製造物責任の保護を受けることが消費者の合理的期待と解すべきか。

※なお、技術の進展に伴いアップデートを行うべき状況が生じることも想定され、その際にCが必要なアップデートを行わなかった場合、過失責任の有無が論点となると考えられる。

## 【論点6-6】立証上の論点

- 想定事例6との関係では、以下のような欠陥の推定や資料が重要と考えられる。

### ➤ Cの責任に関する立証手段

- ✓ 製造物責任訴訟では、裁判例上、「通常の用法に従って使用していたにも関わらず、身体・財産に被害を及ぼす異常が発生した」ことをもって欠陥の主張立証として足りるという欠陥の事実上の推定が活用されている（仙台高判平成22年4月22日（判例集未登載）、東京高判平成25年2月13日判時2208号46頁）<sup>18</sup>。自律的に稼働するロボットについては以下の点が論点となる
  - ・ 「通常の用法」とはどのようなものか：重要な点において想定利用環境や利用方法を逸脱していない状態がこれに該当するのではないか。想定事例6では、AMRが合意された想定環境下で稼働し、作業員がDの定めた運用マニュアルや教育指導の内容を逸脱していないこと等がポイントになると考えられる。
  - ・ どのようなケースが「異常」に該当するか：ロボットの挙動のバリエーションが増えるほど、意図した設計による挙動なのか、「異常」な挙動なのかの判断が困難となる場合も想定される。想定事例6との関係では、突如として危険な走行を始めたような場合は「異常」に該当すると評価されやすいが、一見して異常か否かが不明瞭な場合、意図した設計から逸脱する挙動であることを被害者において立証する必要があり得る<sup>1</sup>。
- ✓ 重要な資料として、製造物責任訴訟では関連する証拠が膨大に亘ることが少なくないが、ソフトウェアで自律的に稼働する側面との関係では、検知に成功していたか、制動は意図した設計どおりに機能したか、回避や停止の判断は合理的か等を検証する観点で以下のような資料が重要になると考えられる。
  - ・ 運行ログ・イベントログ等、速度や停止の機能状況等を示す記録
  - ・ 車載センサ・カメラの記録・映像
  - ・ 事故現場の監視カメラ映像 等

### ➤ Dの責任に関する立証手段

- ✓ 上記資料に加え、以下のような資料が重要になると考えられる。
  - ・ リスクアセスメントの結果を記載した文書
  - ・ リスクアセスメントに基づき設計された現場レイアウトや運用マニュアル等の内容を示す文書 等

18 自動運転車に関して、通常の用法に従って使用していたとしても自動運転装置がすべての事故を防止できるわけではないところ、自動運転装置によって回避できなくてはならない事故であったかが争点となり、この点の立証が必要になるとの指摘について、藤田友敬「自動運転をめぐる民事責任法制の将来像」藤田友敬編『自動運転と法』279頁（2018年、有斐閣）参照。

## 参考裁判例

### 参考裁判例⑫：人が操作するフォークリフトに関する事故（東京地判平成28年2月16日（判例集未登載））

- ・ フォークリフトの運転に伴って生じた人身事故の事案で、運転手には「本件車両を運転して本件駐車場内を後退するに当たり、後方の安全を確認して進行する注意義務を怠り、漫然と後退した過失がある」と認定した上で、当該運転手の使用者である企業に使用者責任を認めた。
- ・ その上で、被害者は、フォークリフト等の車両が往来する事故現場において、本件駐車場内を歩行するに当たって車両等の通行を予見し進路の安全を確認する義務を怠り、本件車両が回転灯を点灯しバックブザーを鳴らして右方から接近してきたにもかかわらず、漫然と歩行した過失があるとして、原告と運転手の過失割合を1対9と認定した。

➢ 従来、フォークリフト等の人が操作する機械により生じた事故においては、運転手の過失責任を認めた上で、事業場を運営する企業に対し使用者責任を問うことにより被害者の救済を図る解決が採られてきた<sup>19</sup>。

### 参考裁判例⑬：チャイルドシートの欠陥（広島地三次支判平成19年2月19日（判例集未登載））

- ・ 交通事故に際し、子に着用させていたシートベルトの肩ベルトが外れ子が投げ出され死亡したため、本件チャイルドシートには瑕疵があったとして、両親が製造会社に対し製造物責任に基づく損害賠償を求めた事案
  - ・ 裁判所は、日本工業規格に適合するチャイルドシートは、特段の事情がない限り、一応その拘束性において欠陥のない製品であると推測されるところ、当該推測を覆すに足りる特段の事情が認められない本件では、チャイルドシートに欠陥があるとはいえないと判示した。
- 日本工業規格に適合しているチャイルドシートは、特段の事情がない限り、一応その拘束性において欠陥のない製品であると推測されると判示した裁判例があり、詳細な安全基準を定める規格は欠陥の判断に当たって重要な考慮要素の一つとなり得ると思われる。
- もっとも、JIS規格は工業用品等の品質を標準化することを目的とするから、同規格に合致しなかったとしても欠陥とはならないとした事例（東京地判平成19年11月27日（判例集未登載））もあり、絶対的な基準とはならないため、最終的には個別の事案ごとの検討をする。

19 同様の解決を行った事例として、大阪地判平成25年7月12日（判例集未登載）、東京地判平成18年4月7日労判918号42頁等がある。