経済産業省

資料３

第１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8月19日

商務情報政策局　情報経済課

1．研究会の趣旨・目的及び検討対象

2．想定事例1：判断補助AI（通常業務）

3．想定事例2：判断補助AI（専門業務）

4．複数当事者間の責任範囲に関するフレームワーク―想定事例2を題材に

AIの利活用に伴う民事責任

- AIの普及に伴い、第三者の財産的権利の侵害や、アクチュエータ（駆動装置・作動機構）を通じた物理的な事故等の発生が懸念される中、インシデント発生時の民事責任の所在について検討を進める必要。

現状の課題

①予測可能性の向上

AI利活用に伴う不法行為法・製造物責任法の解釈適用が不明瞭
→利用や開発への萎縮効果

②ガバナンスの実効化

AI事業者ガイドラインと責任論との関係性が明確でない[1]
→ガバナンスが遵守されず、リスクが顕在化する恐れ

③迅速な事故処理

事案の解決に当たり、高度な専門技術的知見が必要
→裁判が長期化し、迅速な事故処理や被害回復が達成されない懸念

検討の方向性

有識者の議論を取り纏めた準則の策定を目指す[2]

- 現行の法令を前提に、AI利活用特有の論点や解釈の方向性を議論
- AI事業者ガイドラインと責任論との関係性も検討
- 関係者に論点の所在及び考え方の指針を提供することで、迅速かつ円滑な事故処理や被害回復に繋げる

1 AI事業者ガイドラインの検討会においても、責任論に関する検討の必要性を指摘する意見が複数寄せられた（総務省・AIネットワーク社会推進会議（第 30 回）AI ガバナンス検討会（第 26 回）、 経済産業省・第4回AI事業者ガイドライン検討会）。
2 参考：「電子商取引及び情報財取引等に関する準則」を改訂しました （METI/経済産業省）

従来の関連ガイドライン及び本研究会の検討対象

- 契約実務やガバナンスの在り方に関する従来の議論も踏まえつつ、不法行為法や製造物責任法に関する解釈適用の在り方について議論を深め、AIの利活用を促進するツールとすることを目的とする。

✓AI・データの利用に関する契約ガイドライン
✓AIの利用・開発に関する契約チェックリスト

責任論との関係性：
① 経済的な損害等に関する責任を契約において分配する際、責任論の解釈適用が不明瞭であることにより、当事者間で責任の所在に関する目線が合わない場面が生じ得る
② 契約の効力は当事者にしか及ばず、契約外の第三者との関係は不法行為法や製造物責任法によって規律される

本研究会の検討対象

✓ 「責任論」の内容としては、**不法行為法**及び**製造物責任法**を中心としつつ、契約上の債務不履行責任等を含む
✓ AIの自律性[1]やブラックボックス性[2]、AI技術特有の経験則、社会に与える便益の最大化等の観点を踏まえ、過失責任における「**過失**」の有無や製造物責任における「**欠陥**」の有無をどのように判断するか

（※）損害論については、AIの文脈でも従来の理論を援用しうることから、主たる検討対象とはしない

✓AI事業者ガイドライン

責任論との関係性：一般に、安全基準等を定めたガイドラインは責任論の解釈適用に当たっても参照されることがあるが、AI事業者ガイドラインが責任論の観点でどのように評価されるかが整理されていない

1 環境に応じて目的達成に必要な判断や行動を自ら選択し実行する能力を有すること。　2 入力と出力の対応関係や内部の推論過程が人間にとって理解困難または不透明であること。

検討対象ユースケース

- 責任論における解釈適用の在り方は個別の事案ごとに様々であるが、以下のような代表的な想定事例に関する検討を通じ、一定の汎用的な考え方や判断枠組みを抽出したい。

	類型	具体例
①人の行為や判断にAIを活用する類型	自己の行為としてAIに出力をさせる場合	チャットボットAIの出力内容をサービスの内容に組み込むケース
	AIの出力を自己の行為の内容に組み込む場合	画像生成AIが生成した画像を自己の創作として公表するケース
	AIの出力を基に行為を実行する場合	審査業務にAIを活用し迅速化・効率化を図るケース
②AIが自律的に判断を下す類型	AIの判断により経済的な状態が遷移する場合	AIが自ら判断を下して人の代わりに取引を実行するケース
	AIの判断により物理的な状態が遷移する場合	AIが空間認識・機体制御を行って機械を動作させるケース

- 原則としてAIを用いる人が最終的な責任を負う傾向
- ただし例外的な場合には、AIの開発者・提供者等が責任を負う場面があり得る

- 相対的に、AIの設計・開発段階における適切性が問われやすい傾向

不法行為法の概要―過失責任

- 第1回～第2回研究会で検討対象とする想定事例は、基本的に不法行為法のうち過失責任の枠組みで処理される。
- AIサービスやAIシステムとの関係で論点となりやすい責任原因の概要は以下のとおり。

		要件	概要	
一般不法行為（民法709条）		①故意又は過失	行為者が損害の発生を予見し得たか（**予見可能性**）、それを回避する義務が存在したか（**結果回避義務**）	行為者（ここではAI開発者、提供者、利用者等）が被害者に対し、損害を賠償する義務を負う
		②保護法益の侵害	身体・財産・人格権等、法的保護に値する利益の侵害が認められるか	
		③損害の発生	被害者に損害が生じているか	
		④因果関係	①と③の間に事実的因果関係（「あれなければこれなし」）及び社会通念上の相当因果関係があるか	
共同不法行為（民法719条）	**1項前段※**	①複数人の不法行為	複数人の行為がそれぞれ一般不法行為の①～④の要件を満たすか	複数人が連帯して損害の全てを賠償する義務（連帯債務）を負い、被害者は各行為者に損害の全額を請求可能
		②関連共同性	複数人の行為が共同して一つの損害を発生させたか	
	2項	①他人の不法行為	他人の行為が一般不法行為の①～④の要件を満たすか	
		②教唆・幇助行為	他人に不法行為を決意させる積極的働きかけ（教唆）、又は不法行為を容易にする物理的・精神的援助（幇助）	

※ さしあたり1項後段（加害者不明のケース）は省略している。なお、719条2項の責任は1項の責任の一部を注意的に規定したものであり、両者を区別する実益は無いとする見解も有力である。

1．研究会の趣旨・目的及び検討対象

2．想定事例1：判断補助AI（通常業務）

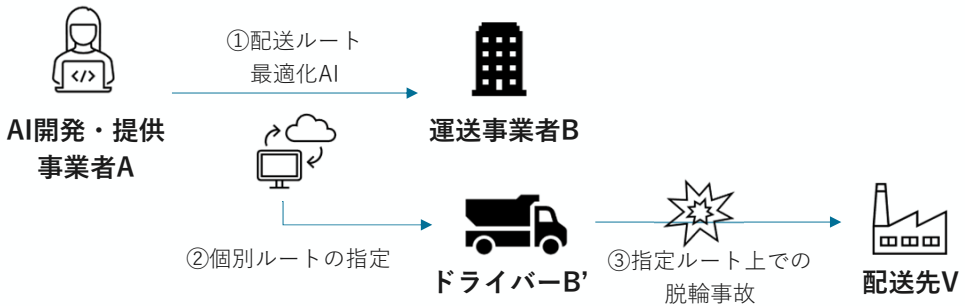
3．想定事例2：判断補助AI（専門業務）

4．複数当事者間の責任範囲に関するフレームワーク―想定事例2を題材に

想定事例1：判断補助AI（通常業務）

- 運送事業者Bは、AIの開発・提供事業を営むAが提供する配送ルート最適化AIシステムを使用し、日々の配送ルートを決定している。当該システムでは、配送先や車両情報、時間制限などの条件を入力することで、配送先までの距離・納品時刻・交通状況等の諸条件を考慮し、最も効率的な配車及びルーティングの計画を自動で作成することができ、配送時間の短縮や必要配車台数の削減等のコスト削減が可能となる。Bは各ドライバーに対し、原則としてAIが出力したルートに従うよう指示していたが、当該システムは幅員や道路構造上の制約を完全には反映できない仕様であるため、実際の道路状況の安全性（安全に通行できる幅・路面・視界が確保されているかなど）は現場にて確認し、安全性が確保できない場合には自ら判断した安全なルートを走行するよう教育・指導を行っていた（当該システムの取扱説明書等にも同様の警告が記載されていた）。

- ある日、BのドライバーB’が運送業務を行っていたところ、幅員が狭く大型車両の運行に適さない悪路を「最適ルート」と表示した。B’は当該道路への進入に躊躇したが、当該日の配車スケジュールが極めてタイトであったことから、全体配送計画との連動性・遅延ペナルティ等を考慮し、AIの出力に従ったところ、結果として運行のための十分なスペースがなく脱輪してしまった。これにより大幅な遅配が生ずると共に、その際の衝撃で配送先V社への荷物が損壊し、Vは代替品の調達等のための損害を被った。



【論点1-1】B及びB’の責任

- 業務上の判断を補助するためにAIを用いた場合、運送会社B及びドライバーB’は、配送先Vに対しどのような内容の注意義務を負うか。AIを用いていない場合における判断と比較して違いはあるか。

➢ B’は荷物を損壊しないよう安全に走行する注意義務を負い、当該義務の違反が認められる場合にはBも使用者責任（民法715条）を負うと考えられる。その際、AIの出力に依拠したことの合理性が論点となる。

参考裁判例：カーナビのルート案内に従って運行した際に車体の擦過傷を生じた事例（福島地判平成30年12月4日判時2411号78頁）

- 原告が、被告らの製造したカーナビシステムのルート案内に従って走行したところ、樹木がせり出した狭い道路に進入し、車体に擦過傷を生じたことから、被告らに対し不法行為及び製造物責任に基づく損害賠償を請求した事案。
- 裁判所は以下の事情を考慮し、「原告は本件カーナビのルート案内に依存せず、自らの判断に基づき本県道路を走行しなければならない」として、ルート案内と擦過傷との因果関係を否定し、被告らの責任を認めなかった。
  ①全国の道路の正確な状況をリアルタイムで情報提供するのは不可能か著しく困難であること
  ②個々の道路の安全性については現に直面する運転者が最も把握し得ること
  ③カーナビの画面や取扱説明書において、実際の道路状況等に従って走行すべき警告等を行っていること
  ④事故現場の道路は国土地理院の地形図に「軽車道」として掲載されており、およそ車両の通行できない道路を収録していたものではないこと

➢ 上記裁判例の論拠（特に①、②、③）は想定事例1にも妥当する。したがって、AIシステムはあくまで**運行上の判断の補助ツール**と位置づけられ、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考えられる。

➢ 他方、以下のような場合には、B’がAIの出力に依拠したことが不合理とはいえず、B’以外の注意義務が論点となる場合があり得る。

- AIの出力に沿わない運行が多発すると全社的な配送計画が乱れることから、Bが各ドライバーに対し、AIの判断に必ず従うよう義務づけていた場合⇒Bの体制構築上の注意義務が論点となり得る。
- AがAIシステムについて重要な説明を行わず、BないしB’が適切な判断ができなかった場合⇒Aの注意義務が論点となり得る（後記10頁）

【論点1-2】Aの責任

- AIシステムを提供していたAは、事故との関係で、配送先Vに対し何らかの注意義務を負うか。

➢ 以下の考慮を踏まえると、Aが何らかの注意義務を負う場面は限定的と考えられる。

- Aは、契約の相手方であるBとの関係ではAIの品質や性能に関する契約上の義務を負うものの、このような契約上の義務は契約外の第三者に対して負うものではない
- 最終的に判断を下したAI利用者（ここではB'ないしB）の行為について責任を負わないのが原則
- 【論点1-1】のとおり、AI利用者にはAIを用いていない場合と同水準の義務が認められることから、重ねてAに責任を認めずとも被害者救済に悖ることはない

→第三者との関係では、原則としてAIに基づき判断を行った利用者が責任を負い、Aの責任が問題となる場面は例外的。

※但し、Aが不法行為法上責任を負わないとしても、契約で合意された品質に満たないAIサービスの提供によって事故を生じた場合には、契約責任を負い得る。

➢ ただし、例外的にBとの共同不法行為責任が問題となり得る場面として、以下が考えられる※。

類型	要素	参考裁判例	想定事例
重要な事項についての説明が欠けたことによる責任（説明義務違反）	① AIへの信頼の程度 ② AIリスクの質・程度 ③ 情報の偏在の程度 ④ 利用者の専門性　等	• 医療機器についての重要な事項に関する説明の欠缺：東京高判H14.2.7判タ1136号208頁等	AIが車両情報や道路状況も踏まえ安全なルートを選択するかのような外観を有していたが、実際のアルゴリズムはそのような計算を行わず、目的地の位置関係や所要時間だけを踏まえルート選択をしており、Aがそのことを説明しなかった場合
侵害行為を容易にしたことによる責任	① AIリスクの質・程度 ② AIが利用者の行為を規律する程度 ③ 結果回避措置の容易性　等	• 著作権に対する寄与侵害：最判H13.3.2最高裁HP・東京高判H11.11.29最高裁HP • 詐欺に用いられた銀行口座の開設：東京地判H28.3.23　等	Aは、テスト等を通じてAIが大型車を狭い幅員の道路に誘導してしまう傾向があること、運送会社の指示等によりドライバーがAIの出力に拘束されがちであること等を把握しており、同種トラブルの報告が多数寄せられていたにもかかわらず、設計の変更や注意喚起等の措置を講じていなかった場合

※ 理論上はAが悪意で権利侵害を引き起こしたような場合にも責任が生じ得るが、そのような事案は極めて例外的と思われる。

1．研究会の趣旨・目的及び検討対象

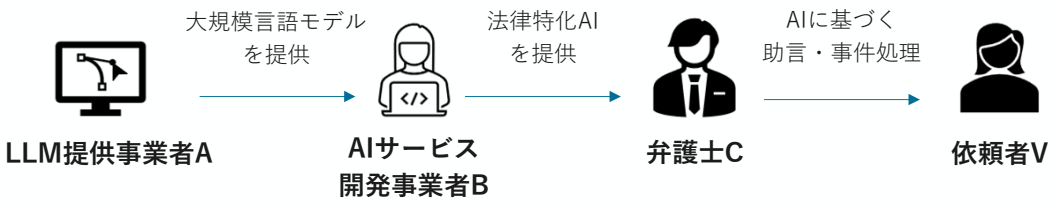
2．想定事例1：判断補助AI（通常業務）

3．想定事例2：判断補助AI（専門業務）

4．複数当事者間の責任範囲に関するフレームワーク―想定事例2を題材に

想定事例2：判断補助AI（専門業務）

- LLM提供事業者Aは、自社が開発した汎用的な大規模言語モデルの提供を行っている。AIサービス開発事業者Bは、法的な推論能力を向上させる目的でAのモデルをファインチューニングした上、推論時に自社で作成した法令・裁判例データベースから関連情報を検索してモデル入力に統合する（Retrieval-Augmented Generation, “RAG”）ことにより、弁護士業務を補助するチャットボット形式のAIサービスを構築・提供している。当該AIサービスでは、弁護士が調査したい内容や具体的な事案をプロンプトとして入力することにより、関連性の高い文献や裁判例を表示したり、法的な分析を網羅的に纏めたレポートを生成する機能を有している。弁護士Cは、当該AIサービスを自己の業務に導入している。
- Cは、依頼者Vから相談を受けた紛争案件について、従来の経緯書や証拠書類をAIサービスに入力し、当該案件に対する法的論点や事件の見立てに関する分析を求めた。AIの出力によれば、Vの立場を裏付ける裁判例が複数存在し、主張が認められる可能性は高いという結論であった。このような分析結果はC自身の経験や考え方とも合致するものであったため、CはAIの出力を信頼し、和解を検討せず、Vの利益の最大化を図る方針を立てた。しかし実際には、当該AIは実在しない架空の裁判例を複数引用し、Vの立場を支える根拠としており、実存する裁判例としてはVの主張に否定的なものが大多数を占めていた。
- 最終的に、裁判所においてVの主張は全面的に棄却され、Vは訴訟追行等に要した手続費用等の損害を被った。



【論点2-1】Cの責任

- 業務上の判断を補助するためにAIを用いた場合、弁護士Cは、依頼者Vに対しどのような内容の注意義務を負うか。
- 想定事例1と比較して違いはあるか。

➢ Cの責任：Vとの間の委任契約上、事件処理にあたり重大な過誤によってVに経済的損失を生ぜしめないようにする注意義務を負う。

参考裁判例：弁護士の顧客に対する助言の過誤に関する想定事例

- 東京地判平成8年4月15日判時1583号75頁：店舗賃貸借契約に関する紛争処理において、弁護士が依頼者から賃料の預託を受けたにもかかわらず、貸主に対する支払いや供託を行わなかったところ、それが原因で賃貸借契約が解除されたことが委任契約の債務不履行に当たるとされた。
- 東京地判平成28年8月24日判タ1433号211頁：相続に関する紛争処理において、被相続人から孫に対して不動産の所有権移転登記をした相続人が単純承認したものとみなされたことにつき、弁護士に説明義務違反があるとされた。

- 補助ツールであるAIの判断を鵜呑みにせず、利用者が正しい判断を下すべきこと（前記9頁参照）は本想定事例にも同様に当てはまる。
- 加えて、弁護士法72条では、弁護士以外の者が一定の法律事務を行ってはならない旨が規定されており、法律事務に関する判断は弁護士の職責においてのみ行うべきことが前提となっている。こうした規制に代表されるように、専門性の高い業務では、AIを利用する者が自らの独立した判断を下すことが求められている。

【論点2-2】A及びBの責任

- AIサービスを提供していたA及びBは、Cの依頼者Vとの関係で何らかの注意義務を負うか。
- 想定事例1と比較して違いはあるか。

➢ 原則としてAI利用者の行為についての責任を負わないこと（前記10頁参照）は本想定事例にも同様に当てはまる。したがって、想定事例1と同様に例外的なケースでのみ責任が生じ得るが、以下のような違いが考えられる。

- 専門性の高い業務で用いられるAIは、上述のとおり、最終的に利用者である専門家が独立した正確な判断を下すべきことが前提となる。
- 他方、専門家の独立した判断を歪めるような説明義務違反が問題となり得るほか（後記16頁以下参照）、AIのサービス提供による幇助責任が成立する余地もあるが、その範囲は通常業務と比較すれば狭まり得る。

e.g.) 専門業務領域におけるAIの出力は、最終的に専門家による慎重なレビューを介して利用されることが前提となっている。
また、AIを利用する専門家は、自己の業務に適したツールを選択する義務を含めた高水準の注意義務を負うと考えられる。
⇒専門家による誤った判断をAIサービスの誤出力やAIサービスに関する説明内容に帰責しうる余地は限定的ではないか。

➢ 仮にAIサービス提供者の責任範囲が異なり得る場合、どのような業務が「専門業務」に該当するか。

- 両者の区分はあくまで相対的なものであり、明確な区分は存在しないが、以下のような事情を考慮し、人がAIの出力にかかわらず独立した正確な判断を下すことが求められる業務は、「専門業務」と評価される可能性が高まると考えられる。

① 業務の性質・難易度
② 判断に誤りがあった場合のリスクの性質・程度
③ 資格制度の有無

1．研究会の趣旨・目的及び検討対象

2．想定事例1：判断補助AI（通常業務）

3．想定事例2：判断補助AI（専門業務）

4．複数当事者間の責任範囲に関するフレームワーク―想定事例2を題材に

複数当事者間の責任範囲に関するフレームワーク―想定事例2を題材に

- 想定事例2において、Vから以下のような説明義務違反が主張されたと仮定する。想定事例2のAIサービスはA・Bの関与を跨いで提供されているところ、それぞれどの範囲で、どのような説明義務を負い得るか。

➢ AIに内在する不確実性に対処するためには、関係当事者間においてAIの性能限界等に関する情報を連携することが重要であり、民事責任論上も、このような情報伝達を行っていなかった場合における説明義務違反（又は指示・警告上の過失）が重要な論点となり得る。
➢ 想定事例2との関係では、以下のような説例が想定される。

- 説例(i)：RAGのデータベースには特定の法分野に関するデータが存在しておらず、当該法分野においては不正確な出力が頻発しやすい状態となっていたところ、Bはサービス全体の出力精度が高いことをアピールする一方、当該リスクを全く説明していなかった。
- 説例(ii)：Aの大規模言語モデルは、特定の法分野について過度に強化学習（RLHF）を行うと、当該法分野については回答精度が向上するが、当該法分野での考え方を別の法分野にも誤って援用し、架空情報を生成する可能性が著しく高まる状態にあった。そのことが説明されていなかった結果、Bが過度な強化学習を行い、特定の法分野において回答精度が著しく低い状態となった。

ポイント	概要
説明義務の認定に当たり一般に重要となる要素	• 前記10頁のとおり、AIリスクの質・程度、情報の偏在の程度、利用者の技能・専門性等が重要な考慮要素となる。
AI事業者ガイドラインに基づく要請	• AI事業者ガイドラインにおける透明性の要請も説明義務を基礎づける一事情となりうる（後記17頁「**AI事業者ガイドラインにおける透明性の要請**」参照）。
バリューチェーンにおける各当事者の役割	• 説明義務を含む過失の認定に当たっては、結果回避可能性（結果回避措置を取り得る立場にあったこと）が不可欠となる。このような観点から、各当事者は原則として、AIのバリューチェーンにおける様々な要素のうち、自ら主体的に関与して、仕様や挙動をコントロールした事項についてのみ責任を負う。 • 想定事例2においては、Aは基盤となる汎用モデルのチューニングサービスを提供する立場である一方、Bがモデルの挙動調整や詳細なサービスの設計など、広範な要素に関与している（後記18頁「**AIサービスのバリューチェーン**」参照）。

AI事業者ガイドラインにおける透明性の要請

- AI事業者ガイドラインでは、総論としての「共通の指針」及び開発者に関する各論としての「AI開発者に関する事項」において以下の透明性確保措置を求めており、こうした行動規範も説明義務を基礎づける一事情となり得る。

全事業者が遵守すべき「共通の指針」[1]

② 関連するステークホルダーへの情報提供

- ✧ AIとの関係の仕方、AIの性質、目的等に照らして、それぞれが有する知識及び能力に応じ、例えば、以下について取りまとめた情報の提供及び説明を行う
  - ● AIシステム・サービス全般
    - ➢ AIを利用しているという事実及び活用している範囲
    - ➢ データ収集及びアノテーションの手法
    - ➢ 学習及び評価の手法
    - ➢ 基盤としているAIモデルに関する情報
    - ➢ AIシステム・サービスの能力、限界及び提供先における適正/不適正な利用方法
    - ➢ AIシステム・サービスの提供先、AI利用者が所在する国・地域等において適用される関連法令等
- ✧ 多様なステークホルダーとの対話を通じて積極的な関与を促し、社会的な影響及び安全性に関する様々な意見を収集する
- ✧ 加えて、実態に即して、AIシステム・サービスを提供・利用することの優位性、それに伴うリスク等を関連するステークホルダーに示す

「AI開発者に関する事項」[2]

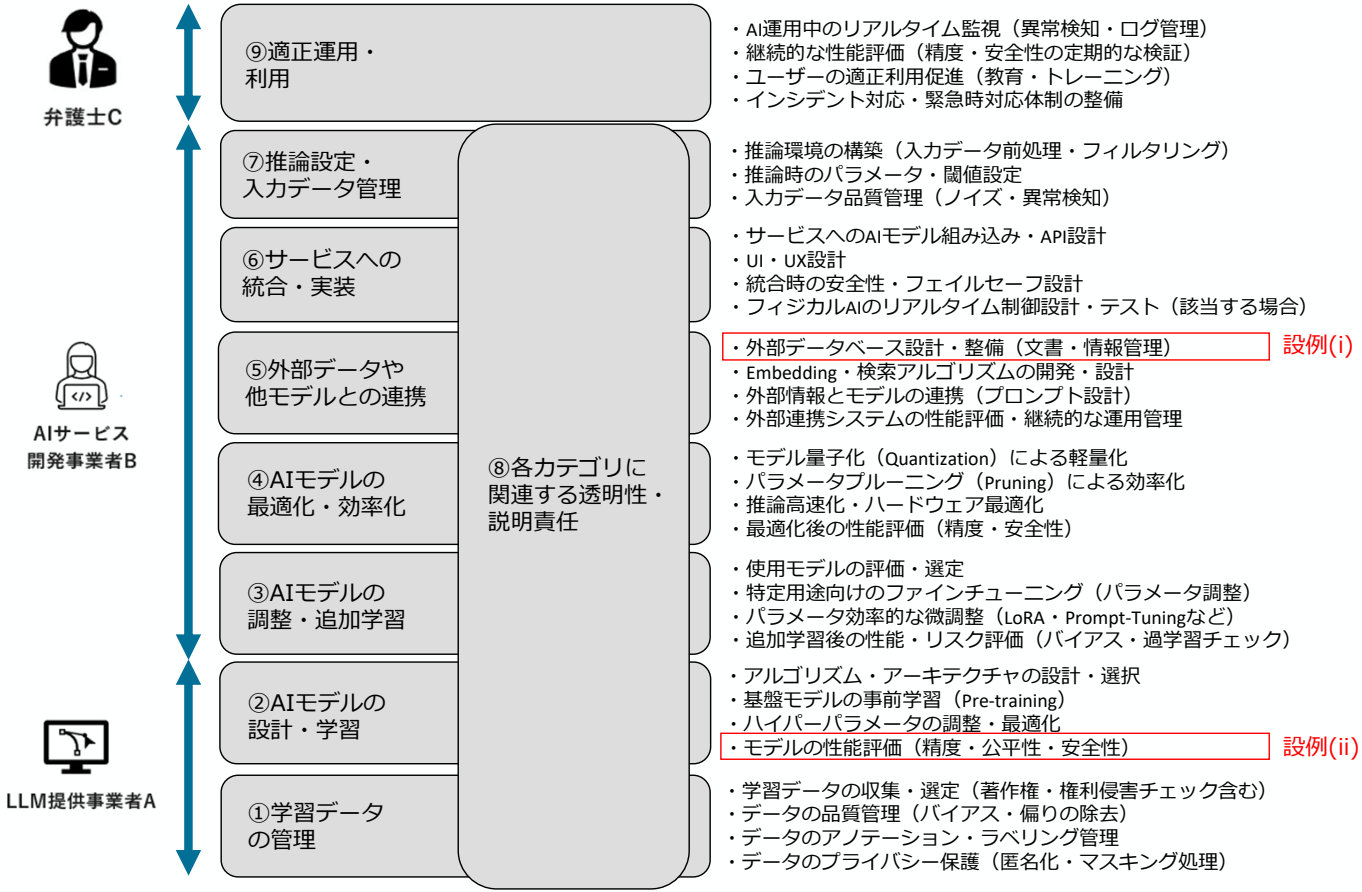
➢ D-6）ii. 関連するステークホルダーへの情報提供

- ✧ 自らの開発するAIシステムについて、例えば以下の事項を適時かつ適切に関連するステークホルダーに（AI提供者を通じて行う場合を含む）情報を提供する（「6）透明性」）
  - ● AIシステムの学習等による出力又はプログラムの変化の可能性（「1）人間中心」）
  - ● AIシステムの技術的特性、安全性確保の仕組み、利用の結果生じる可能性のある予見可能なリスク及びその緩和策等の安全性に関する情報（「2）安全性」）
  - ● 開発時に想定していないAIの提供・利用により危害が発生することを避けるためのAI開発者が意図する利用範囲（「2）安全性」）
  - ● AIシステムの動作状況に関する情報並びに不具合の原因及び対応状況（「2）安全性」）
  - ● AIの更新を行った場合の内容及びその理由の情報（「2）安全性」）
  - ● AIモデルで学習するデータの収集ポリシー、学習方法及び実施体制等（「3）公平性」、「4）プライバシー保護」、「5）セキュリティ確保」）

バリューチェーンの観点で主体間の連携を確保し、全体としてリスクコントロールを図る
というAIガバナンスの在り方[3]との関係で、中核的な要素

1 総務省・経産省「AI事業者ガイドライン（第1.1版）」（2025年3月28日）本編19頁 https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_1.pdf
2 同31-32頁　3 同28頁

AIサービスのバリューチェーン



A及びBが負いうる説明義務

- 前頁までの考慮事項を総合すると、Vの主張する説明義務違反が認められる余地はあり得るか。

ポイント	想定事例2における考慮要素
説明義務の認定に当たり一般に重要となる要素（AIリスクの質・程度、情報の偏在の程度、利用者の技能・専門性等）	• 利用者である弁護士は、AIの品質にかかわらず、法律の専門家の立場からAIの出力を独立に吟味検証すべき立場にあるため、一般的なリスクまで含めて広く説明義務の対象となるものではない • 他方、技術的・潜在的なリスクであって、利用者による誤った利用やAIに対する信頼の程度の誤認を誘引するようなリスクは説明義務の対象となり得る（前記9頁の参考裁判例も参照） • **評価が分かれ得る点：架空の裁判例を挙げるというハルシネーションは、単に誤った情報を生成する場合と比べ、一見して生成内容に対する「もっともらしさ」を生じやすい等の観点で、法的評価に差はあるか。**
AI事業者ガイドラインに基づく要請	• AIサービスの開発に関与する者は、AIの「能力」や「限界」、想定される「適正／不適正な利用方法」等に関する適切な情報開示が求められ、説例(i)及び(ii)はこうした開示要請に合致する • 透明性確保措置は、バリューチェーン全体でのリスク管理のため重要な要素
バリューチェーンにおける各当事者の役割	• Aは汎用的な言語モデルを提供するに留まる一方、Bがシステムやサービスの構築全般に亘る広い要素に関与している。説例①はBの関与範囲に属し、説例②は基本的にAの関与範囲に属する。

- 説例①（特定の法領域における性能劣化）

  ：RAGのデータベース構築を含むサービス全般の設計を行ったBに説明義務違反が認められる可能性があり得るか。

- 説例②（過度な強化学習による著しいハルシネーションリスクの増大）

  ：基盤モデルを開発し提供していたAに、Bに対して必要な情報を提供しなかった説明義務違反が認められる可能性があり得るか。

  - ただし、一般的なAI開発者であれば検知できたであろう程度のリスクの場合、Bにおいて当該リスクを把握しリスク軽減策を講じるか、又はBからCに対してリスク告知すべきであったという評価も考えられ、Aはそのようなリスクについて説明義務を負わない、あるいは説明義務違反と損害発生との間に因果関係がないと判断され得る。

重要な証拠

- Vの請求原因や、A及びBからの反証の観点では、例えば下表のような資料が重要な証拠となる。
- 一部の資料は任意開示や文書提出命令の対象となり得るものがあるが、開示可否に争いが生じることが予想されるものについては可能な限り契約等で取扱いを明確にしておくことが望ましい。

	証拠	立証命題
Cから提出しうる資料	API入出力ログ（質問（prompt）、回答（completion）、時刻（timestamp）など）	架空の裁判例等の誤情報が実際に出力された事実
	基盤モデルに関するModel Card／System Card、マーケティング資料	基盤モデルに関し公表されている性能や概要
	Bのサービスに関する性能説明書やリスク説明書	サービスの性能や品質に関して顧客に提供された情報
	外部ベンチマーク評価 ※統計的に有意な量のサンプルが用意できる場合	サービスの精度を客観的な指標を用いて評価した結果
Bから提出し得る資料	強化学習（RLHF）済みモデルやシステムに関するリスク分析報告書	B自身が提供サービスについて分析したリスクの内容
	RAG検索ログ（search_query, hit_count, top_k, score等）	プロンプトの処理経過（データベース内にヒットする文書が存在しなかったこと等）
Aから提出しうる資料	基盤モデルに関するリスク評価報告書（一般的ハルシネーション、RLHFによる影響、性能限界等）	A自身が基盤モデルについて分析したリスクの内容

➢ 上記のほか、A及びBにおけるAIガバナンスの状況を客観的に示す資料として、各主体が運用するAIシステムの開発に関するマネジメントシステム（例：ISO 9001、ISO/IEC 42001等の要素を含むもの）に基づき適切に維持・管理されるべき文書化された情報も、重要な証拠資料となり得る。