

AI 保険の法的課題と展望

—海外動向の分析を中心に—

東京大学大学院法学政治学研究科 柳平大樹

要旨

生成 AI の社会実装に伴い、ハルシネーション、アルゴリズムバイアス、サイバー攻撃の高度化、機密データ漏洩等のリスクが顕在化している。本稿はサイバー保険の発展過程を手がかりに、AI 保険の法的課題を検討する。

サイバー保険の判例

Merck 判決・G&G Oil 判決を分析し、約款精緻化の教訓を抽出する。

海外市場の二極化

AI 除外条項と AI 専用保険 (Armillia AI/Chaucer、Munich Re) への分化を検討する。

日本法への含意

告知義務、AI 除外条項の解釈、AI ガバナンス体制構築義務、保険による規律付けを論じる。

❏ **中心的な問い:** AI 保険は「保険による規律付け」として機能し得るか。そのためにいかなる基盤が必要か。

キーワード

本稿の主要概念

AI 保険

AI システムの開発・運用・利用に起因するリスクを填補対象とする保険の総称。専用保険とサイレント AI 問題の二形態を包含する。

サイバー保険

AI 保険の前身。リスク構造の類似性と実務上の連続性において AI 保険と特別な関係にある。

保険による規律付け

保険会社がリスク評価や引受条件を通じて被保険者の行動に影響を与え、社会的に望ましい行動を促進する機能。

AI ガバナンス

内部統制システム構築義務の一内容として取締役に求められる AI ガバナンス体制構築義務。

はじめに—問題の所在

AI インシデントの急増

Stanford University の調査によると、AI インシデントの件数は 2012 年以降 26 倍に増加している。意思決定者の 2 人に 1 人が AI のリスクはメリットより高いと考えているとの報告もある (Munich Re)。

日本市場の現状

日本の AI 保険市場は黎明期にあり、保険法からの AI リスク分析も始まったばかりである。産業界は AI の社内活用や AI ガバナンス構築という課題に急ピッチで取り組んでいる段階であるといえ、本格的な保険の検討までは至っていないものともいえそうである。

三つの課題

1. サイバー保険の発展過程から AI 保険の約款設計への教訓
2. 海外における AI 保険の動向が日本に持つ含意
3. AI 保険と AI ガバナンス体制構築義務の関係

AI リスクの類型化

ハルシネーション

AI 固有の最顕在化リスク。生成 AI は確率論的に誤情報を出力し、存在しない商品案内や法的に誤った助言が現実には発生している。EIOPA 調査では EU 保険会社が認識する最大リスクとされる。

サイバー攻撃の高度化

生成 AI により巧妙なフィッシングメールやマルウェアが作成可能となり、サイバー攻撃の脅威が増大している。

アルゴリズムバイアス

AI モデルが学習データの偏りを反映し、特定の人種・性別・年齢層に不利な判断を下す。採用・融資審査・保険引受等で差別訴訟リスクとして現実化している。

ブラックボックス性・集積リスク

深層学習モデルの判断過程は理解困難であり、因果関係立証の障害となる。少数の基盤モデルへの依存はシステムック・リスクを内包する。

従来型保険の限界

企業はデジタルリスクへの備えとしてサイバー保険を活用してきたが、AI 特有のリスク——とりわけ AI の自律的判断や出力内容の瑕疵に起因する損害——は従来のサイバー保険や賠償責任保険でカバーするのが困難なケースが多い。

従来型保険

外部サイバー攻撃・不正アクセスを主に補償するが、AI固有のリスクはカバーされていない。

AI保険

ハルシネーション、性能劣化、自律的判断の瑕疵に起因する損害を明示的に補償する。

サイバー保険と AI 保険の間には構造的・実務的な連続性がある。不確実性の高い新技術リスクである点、システムック・リスクの可能性を有する点、外部攻撃・内部瑕疵・利用者行為の複合によって損害が生じる点で両者は共通の構造を持つ。

サイバー保険からの教訓

海外ではサイバー保険は約款の曖昧さから多くの紛争を経験し、裁判例を通じて解釈が精緻化されてきた。保険による規律付けが機能するには、まず約款がAIに関する実態に即して精緻化されている必要がある。

市場の概況

米国等では加入率が上昇する一方、損害率の上昇により保険料率の引上げや引受条件の厳格化が進んでいる。米国元受保険市場における保険料は2021年に約75%、2022年に約50%増加した(NAIC)。

PL保険・CGL保険との比較

新技術リスクに起因する保険約款の精緻化という現象はPL保険・CGL保険においても見られる。しかし、サイバー保険はリスク構造の類似性および実務上の連続性においてAI保険と特別な関係にある。

Merck 判決—戦争免責条項の再定義

Merck & Co. v. Ace American Insurance Co. (N.J. Super. Ct. App. Div. 2023)

事案の概要

NotPetya マルウェア攻撃 (2017 年) による約 14 億ドルの損害について、全危険型財産保険の戦争免責条項の適用が争われた。ニュージャージー州控訴審は Merck の請求を認容したが、2024 年 1 月に州最高裁での口頭弁論直前に和解で終結した。

AI 保険への示唆

裁判所は戦争の概念をサイバー空間に無限定に拡大すべきではないと判示した。保険会社側が不利な先例の確定を回避したとみることができる。LMA はその後、国家関与型サイバー攻撃を明示的に免責とするモデル条項を策定した。曖昧な概念に依存せず、免責範囲を具体的に規定する方向への転換を促した。

G&G Oil 判決—コンピュータ詐欺条項の射程

G&G Oil Co. of Indiana, Inc. v. Continental Western Insurance Co. (Ind. 2021)

事案の概要

ランサムウェア攻撃による身代金支払いが、商業犯罪保険のコンピュータ詐欺条項における「コンピュータの使用により詐欺的に金銭の移転を引き起こしたことに直接起因する損失」に該当するかが争われた。

判旨と意義

インディアナ州最高裁は「詐欺的に移転を引き起こす」を広く解し、FBI への相談等の介在行為があっても因果的近接性は維持されるとして差し戻した。サイレント・サイバー問題の典型例として、従来型商業犯罪保険がランサムウェア損害をカバーし得ることを示した。

サイバー保険判例の小括—AI 保険への教訓

1

約款の曖昧さ

従来の保険約款が現代の複雑なデジタルリスクに対応しきれないことが顕在化した。

2

具体的・限定的設計へ

Merck 事件・G&G Oil 判決はいずれも、具体的に限定的な約款設計への移行を促した。

3

AI 保険への応用

AI 起因の意味を明確に定義し、免責範囲を具体化する必要がある。リスクの具体化と約款の精緻化というプロセスは AI 保険も同様にたどることになるものとする。

海外における AI 保険の動向——二極化

AI リスクの不確実性を前に、海外の保険市場は除外と専用化という二方向へ分化しつつある。この二極化は同時に、リスク評価の困難、因果関係立証の複雑さ、保険と企業ガバナンスの関係といった AI 保険固有の課題を顕在化させる。

AI除外条項

- ・ 既存保険からAIリスクを排除
- ・ サイレントAI問題の回避
- ・ 補償の空白が生じるリスク

AI専用保険

- ・ AIリスクを積極的に引受
- ・ デューデリジェンスによる引受審査
- ・ ハルシネーション・性能劣化を明示補償

AI 除外条項とサイレント AI 問題

米国では複数の保険大手グループが、AI ツールを導入する企業関連で生じる賠償責任を免責とする許可を規制当局に申請している (Financial Times, 2025)。

定義の曖昧さ

初期の AI 除外条項は定義が曖昧であり、原因関係の入口文言が広く、過剰に包括的である。

除外範囲の過大さ

AI が周辺的に使用されているに過ぎない事故まで免責となりかねない。例えば、マーケティング資料作成に AI を使用していたという周辺事情を根拠に専門職賠償請求が免責されるリスクがある。

サイレント AI

約款の曖昧さにより、保険会社が意図せず AI リスクを引き受けてしまっている、あるいはその逆の状態。サイレント・サイバー問題と同様の構造である。

AI 専用保険の開発—Armilla AI / Chaucer

Affirmative AI Liability Insurance の特徴

- ・ ハルシネーション・モデルドリフト等による損失を明示的に補償
- ・ AI ツールの精度が事前設定基準を下回った場合を補償トリガーとするパラメトリック的設計
- ・ 最大 2,500 万ドルの保険限度額 (2026 年 1 月改訂)
- ・ クレームズ・メイド方式

2026 年改訂後の補償範囲

- ・ ハルシネーション・性能劣化による第三者の経済的損失
- ・ 有害出力に起因する名誉毀損・営業秘密露出
- ・ 自律 AI エージェントの誤判断・不適切なツール使用
- ・ AI 出力を通じた意図しない情報漏えい
- ・ EU AI Act 等の規制法に基づく当局調査対応費用・行政制裁金

同社 CEO: *"It's built from the ground up to address the specific ways AI can fail."*

Armilla AI の二層構造——保険と保証

- Layer 1: Armilla Insured

AI導入企業向け賠償責任保険

- Layer 2: Armilla Guaranteed

AIベンダー向けパフォーマンス保証 (Swiss Re・Greenlight Re・Chaucerがバックアップ)

Armilla Guaranteed は AI が契約上の KPI を満たさない場合に買い手への返金を保証するものであるが、同社は「保証は契約上の履行保証であり、保険契約ではない」と明言する。

エア・カナダ判決—ハルシネーションの企業責任

Moffatt v. Air Canada, 2024 BCCRT 149

エア・カナダ社の AI チャットボットが遺族向け割引運賃について旅行完了後にも遡及的に申請可能である旨の誤った案内を行い、顧客がその案内に基づいて航空券を購入した。同社は「チャットボットは別の法的主体でありその発言について責任を負わない」と主張したが、ブリティッシュコロンビア州民事紛争解決審判所(BCCRT)はこれを退け、企業がウェブサイト上の全ての情報について責任を負うと決定した。

- ❑ 本件は BCCRT の決定であり法的拘束力は限定的であるものの、AI エージェントの行為を企業の責任に帰属させた象徴的な事例として実務に衝撃を与えた。Virgin Money・DPD でも同様の問題が発生している。

Munich Re — パフォーマンス保証保険

商品の特徴

技術デューデリジェンスを通じて AI モデルの品質を評価し、その信頼性を保険で裏書きする。クライアントと合意した KPI が補償発動の基準となる。同社はパフォーマンス保証保険を「**信頼の加速装置**」と位置づける。

技術デューデリジェンスを受けた企業の **90%** が保険契約の提供を受けている。

デューデリジェンスの 4 段階

1. 導入ミーティング
2. 詳細技術質問票
3. フォローアップ・ミーティング
4. 報告・結論

保険会社が AI ガバナンスのゲートキーパーとして機能する具体例である。

海外動向の小括—過渡期の位置づけ

二極化の意味

従来型保険では AI リスクは不確実性が高く、サイレント AI 回避のため除外せざるを得ない。他方、デューデリジェンスを経て切り出されたりリスクは専用保険としてプライシングが可能となる。

現在の位置

現在の海外動向は、AI リスクが漠然とカバーされる段階から明示的に管理・移転される段階への過渡期に位置する。

日本への示唆

あいおいニッセイ同和損保が 2024 年 2 月に国内初の生成 AI 専用保険を公表したが、AI 開発会社を保険契約者とする仕組みであり、一般企業が広く加入できる汎用商品ではない。

AI 保険の確立に向けた課題

1 リスク評価と因果関係立証の困難

AI の挙動は確率論的でブラックボックス性が高く、事故発生確率の予測が困難である。損害が学習データのバイアスによるのか、ファインチューニングの不備によるのか、悪意あるプロンプト注入によるのかの区別は技術的に容易ではない。

2 間接差別と代替変数の問題

AI が膨大な非構造化データからリスクを推論する際、法的に考慮してはならない属性と相関する代替変数を用いて差別的選別を行う危険がある。バイアス監査の仕組みの構築が求められる。

3 因果関係論の限界とパラメトリック保険

従来の保険法理が前提としてきた因果関係論は AI リスクにおいて機能し得るのか。全ての損害について因果関係の立証を求めれば、被保険者保護や保険金支払の迅速性が損なわれる。

保険法の強行規定と AI のブラックボックス性

保険法 31 条 2 項 1 号但書は、告知義務違反があっても当該事実と保険事故との間に因果関係がなければ保険者は免責されないとする**因果関係不存在特則**を定める。この特則は片面的強行規定(保険法 33 条1項)であり、被保険者に不利な方向での変更は認められない。

- **問題の核心:**因果関係不存在特則の立証責任は被保険者側にある。AI のブラックボックス性ゆえに、被保険者がこの立証に成功することは技術的に困難である。例えば、基盤モデルのバージョンについて不告知があった場合、被保険者が当該バージョンの相違と事故との間に因果関係がないことを論証するには、モデルの内部動作を解析する必要があるが、深層学習モデルにおいてその解析は現状では実現困難である。

なお、保険法 36 条 4 号は事業者間の保険契約について強行規定の適用を除外する余地を認めており、当事者自治による契約設計の可能性も検討に値する。

パラメトリック保険——因果関係論の限界を克服する試み

パラメトリック保険とは

実際の損害額ではなく、事前に定めた指標が一定の閾値を超えた場合にあらかじめ合意された金額を支払う保険形態。AI リスクでは、モデル精度が一定 % を下回った場合やシステムダウンタイムが一定時間を超えた場合等の客観的指標に基づく設計が考えられる。

吉澤の分析枠組みの適用

吉澤卓哉(2020)の定額給付型損害保険契約論によれば、被保険利益が存在し、パラメータと損害との間に合理的な相関関係が認められ、**賭博禁止とモラル・ハザード防止が確保される限り**、損害保険契約としての性質を維持し得る。

AI モデルの精度低下は被保険者が容易に操作し得ない客観的指標であり、事故偽装型のモラル・ハザードも想定しにくい。

パラメトリック AI 保険の法的構造

01

被保険利益の存在

AI システムの開発・運用・利用に対する経済的利益。システム停止や精度低下により被保険者が損失を被る関係が認められれば足りる。

03

損害のみなし発生

パラメータの閾値超過をもって損害の発生とみなす。実損填補原則の修正であり、吉澤の定額給付型損害保険契約論による正当化が必要となる。

ただし、パラメータの設定が損害との相関を欠く場合には、利得禁止原則の緩和では正当化が困難であり、保険契約ではなくデリバティブ取引として規制される可能性がある点には留意が必要である。

- 保険法上の位置づけ: パラメトリック AI 保険は損害保険契約(保険法 2 条 6 号)として構成することが可能であるが、パラメータ設計の合理性・透明性の確保が法的有効性の鍵となる。保険業法上の認可・届出要件についても別途検討が必要である。

02

保険事故の発生

事前に合意したパラメータ(モデル精度・システムダウンタイム等)が閾値を超えること。客観的・検証可能な指標であることが要件となる。

04

損害額のみなし算定

事前合意した金額を支払う。利得禁止原則との緊張関係が生じるが、パラメータと損害との合理的相関が認められる限り許容される。

日本法への含意—概観

01

告知義務と AI インベントリ

質問応答義務への転換を踏まえた AI 関連告知事項の設計と、シャドーAI 管理の問題

02

AI 除外条項の解釈

約款作成者不利の原則・合理的期待の保護・定型約款規制との関係

03

取締役の義務—比較法的視座

Langenbucher・中村・Kourabas & Tsang の知見を踏まえた AI 時代の善管注意義務

04

内部統制システム構築義務と AI ガバナンス

会社法 362 条 4 項 6 号に基づく AI ガバナンス体制構築義務の内容

05

保険による規律付けの可能性

「保険＝ガバナンス」理論に基づく AI ガバナンス促進の好循環

告知義務と AI インベントリ

保険法 4 条は危険に関する重要な事項について告知義務を課している。AI 保険では、生成 AI の利用有無・利用目的・AI ガバナンス構築状況が危険評価の重要な要素となる。

質問応答義務の帰結

周知のごとく、保険法の制定により、告知義務は自発的申告義務から質問応答義務へ転換しており、保険者が質問しなかった事項については原則として告知義務違反を問えない。抽象的・包括的な質問の効力も限定的に解される。

保険者が AI の利用の有無を問うのみで、基盤モデルの種類・RAG の利用有無・ファインチューニングの有無等の技術的仕様について具体的な質問を設定しない限り、企業のシャドーAI 利用は告知義務違反を構成しない可能性が高い。

シャドーAI と告知義務

告知義務者の知らない事実は原則として告知義務の対象とならない。重過失による不注意については告知義務違反が成立しうるが、その重過失はほとんど故意に近似する著しい不注意に限定して解釈すべきとされる。

適切な AI ガバナンス体制 (AI インベントリの整備・利用申請・承認プロセス・定期的な監査等) が構築されていれば、シャドーAI は容易に検知できたはずであり、経営陣の不知は重過失として構成される余地が生じる。

AI 除外条項の解釈—英国法理との比較

AI 専用保険の多くがロイズ・オブ・ロンドン市場で引き受けられ英国法を準拠法とすることに鑑みれば、英国法理を比較法的視点から参照することが有益である。

引受危険の外延を画定する条項

AI リスク全般を包括的に排除する AI 除外条項。約款作成者不利の原則の適用はさらに限定され、文言の平明な意味と契約全体の文脈に基づく通常の契約解釈原則が優先する。

事後的な除外としての条項

特定の AI 利用態様のみを免責とする限定的な AI 除外条項。約款作成者不利の原則を含む約款解釈法理が広く機能する余地がある。

- **英国判例法理** (*Burnett or Grant v International Insurance Company of Hanover Ltd [2021] UKSC 12*) : *contra proferentem* 原則は真の曖昧性 (*genuine ambiguity or real doubt*) が存在する場合にのみ適用される。

約款作成者不利の原則と合理的期待の保護

約款作成者不利の原則の限界

解釈によっていずれかの解釈に収斂するのであれば、約款作成者不利の原則の適用はない。同原則は複数の解釈が成り立ち得る曖昧性が存在する場合にのみ補助的に機能し、文言が一義的に明確である場合には適用されない。

合理的期待の保護

合理的期待の原則は米国法独自の法理であるが、日本法においても約款作成者不利の原則・信義則(民法 1 条 2 項)・定型約款の不当条項規制(民法 548 条の 2 第 2 項)を通じて考慮されうるものとする。

AI が単に業務の一環で使用されていたに過ぎない事故まで免責されるとなれば、被保険者の合理的期待を害する。日本法の下では AI 除外条項の適用範囲は、**損害と AI の挙動との間に相当因果関係が認められる範囲等**に限定して解釈すべきと解される。

約款解釈の限界と契約交渉の重要性

約款の解釈で救済を得ることには限界がある。AI 除外条項の文言が一義的に明確である場合、作成者不利の原則は適用されず、被保険者は免責条項の文言どおりの効果を甘受せざるを得ない。

→ 更改時の交渉

米国で最も有効な対抗策として用いられているのは、更改時点での免責の削除や限定を交渉することである。

→ 文言の限定

削除が困難な場合には、AI の定義、文言（「に起因する」を「を直接の原因とする」に限定する等）、カーブバック条項の導入を検討する必要がある。

→ 募集文書における情報提供

消費者契約法 3 条 1 項の透明性原則の観点から、AI 除外条項の内容が契約締結時に明確かつ平易に説明されることが求められる。募集文書における情報提供の充実は事後的な紛争予防の観点からも重要である。

取締役の義務と AI——比較法的視座

Langenbacher (2025)
Ownership and Trust

ブラックボックス AI を取締役会の意思決定に使用してはならないという見解を明確に否定。AI の予測は「既知の未知」であり、取締役会が日常的に処理している不確実性の一つに過ぎない。

中村直人 (2023)
経営判断手続と信頼の原則

生成 AI が出力した誤った情報が意思決定の前提事実となった場合、取締役の善管注意義務違反が問題となる。AI 利用について社内ルールを定めていなければ、AI 生成情報に含まれる誤りのリスクが見過ごされるおそれがある。

Kourabas & Tsang (2025)
Human-in-the-Loop

オーストラリア法を素材として、自動化バイアスの問題を指摘し Human-in-the-Loop アプローチを提唱。適切な人材を AI 監視業務に配置し、AI を利用した意思決定の適切な手続を確保することが求められる。

AI 時代における取締役の義務の核心

意思決定の「所有」を維持すること

AI ベンダー選定の合理性

外部の AI ベンダーやシステムへの依拠に際して、その選定の合理性と出力の批判的検証を怠らないこと。米国デラウェア州会社法 141 条(e)・ドイツ判例法理(「Ision」事件)が参考となる。

人間の関与の制度的確保

運用過程において適切な人間の関与を制度的に確保すること。AI の信頼性が高まった時代には、AI を意思決定に参照しないことが善管注意義務違反を構成するという可能性も示唆される(中村)。

AI ガバナンス体制の構築・維持

信頼の原則の適用基盤たる AI ガバナンス体制を構築・維持すること。ガバナンス体制の不備は信頼の原則の適用を妨げ、取締役の善管注意義務違反のリスクを高める。

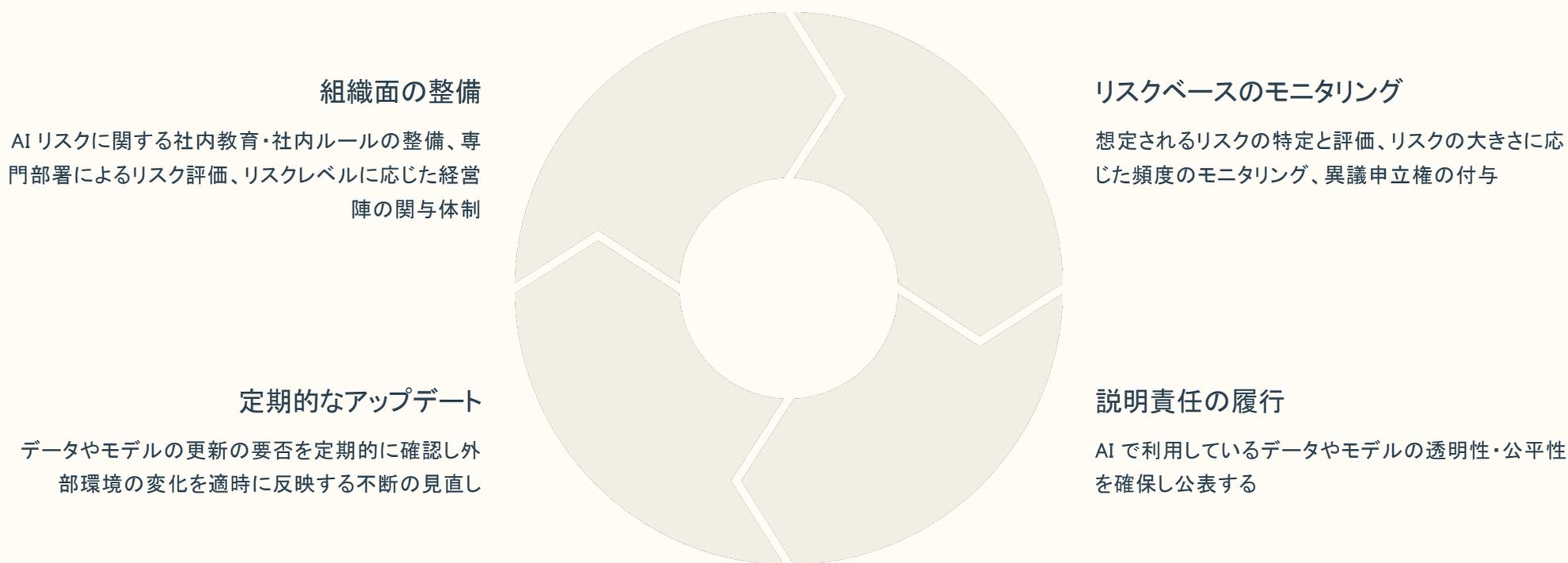
内部統制システム構築義務と AI ガバナンス

内部統制システム構築義務とは、業務の適正を確保するための仕組み・体制を構築する義務である(会社法 362 条 4 項 6 号・5 項、大和銀行事件判決)。AI 活用に伴うリスクは会社法施行規則 100 条 1 項の各号と対応する。

会社法施行規則 100 条 1 項各号	AI リスクとの対応
2 号: 損失の危険の管理に関する体制	AI システムに起因するシステム停止・不正確な出力・情報漏洩等
4 号: 法令・定款適合確保体制	AI の利用が差別的結果を生じさせる場合・個人情報保護法等への抵触
1 号: 情報の保存及び管理に関する体制	AI の判断過程・学習データ等の記録・保存
5 号: 企業集団における業務適正確保体制	子会社等が AI を利用する場合のグループ全体でのガバナンス方針の統一

AI ガバナンス体制構築義務の具体的内容

日本銀行金融研究所報告書(2025年6月)は、AIガバナンス体制構築義務の具体的内容として以下の枠組みを提示する。



- ☐ 同報告書は、AIガバナンス体制の善管注意義務上の水準が静的なものではなく、**技術進化に応じて動的に変化し得ることを示唆する**。グローバルに展開する企業においては海外同業他社の水準を意識することが求められる。

生成 AI 固有の推論段階における技術的統制

EIOPA(2025)は、生成 AI のガバナンスにおいて**推論段階への統制シフト**が生じていることを指摘する。従来の機械学習ではモデル学習段階の検証が中心であったが、生成 AI では基盤モデルが第三者によって事前学習されており、AI ガバナンスの主戦場は推論段階へ移行する。

入力統制

プロンプト入力で何を許容するかを明確にし、プロンプト・インジェクション攻撃や機密情報の意図しない入力を防止する体制の構築(会社法施行規則 100 条 1 項 4 号)

出力監視

ハルシネーションの検出と対処の仕組みの整備。顧客対応に AI を用いる場合には、出力が募集規制上の情報提供義務等の法令に適合するかの確認体制(同 4 号)

権限設計

Agentic AI の導入に際し、外部システムへのアクセス権限の範囲を明確に定義し AI が自律的に行使し得る権限の上限を設定する(同 2 号)

ログ管理

AI の入出力履歴および人間による介入・承認の記録を保存し説明可能性と責任の追跡を担保する体制。保険事故発生時の因果関係立証や規制当局対応においても決定的な意義を有する(同 1 号)

※ 入力統制・出力監視・ログ管理はEIOPA調査30-31頁に依拠。権限設計は筆者の考察による。

保険による規律付けの可能性

「保険＝ガバナンス」理論(王 2024)

サイバー保険の文脈において、保険者は引受決定や契約上の予防措置の要求を通じて被保険者の行動に影響を与え、社会的リスクガバナンスにおいて第三者による監督の機能を果たす。保険者は「準政府 (quasi-government)」ないし「代理政府 (surrogate government)」の性質を有するとも論じられる。

小塚 荘一郎の問題提起

保険商品の前提条件として AI 開発者等に AI 原則の実施や AI ガバナンスの構築を義務づけることで、「AI リスク保険が普及するとともに AI ガバナンスも関係業界に浸透するという好循環が期待できる」(小塚 2021)。

Chaucer は故障に陥りやすい AI の引受けを選別する方針を明示しており、保険会社が AI ガバナンスのゲートキーパーとして機能する実例が既に存在する。

規律付けのエコシステム

- ハードロー

法規制、EU AI Act、NAIC通達など。

- ソフトロー

NIST AI RMF、ISO/IEC 42001などの標準化。

- 市場メカニズム

AI保険、技術デューデリジェンス、料率優遇など。

技術デューデリジェンスを通じて企業の AI ガバナンスを点検し、リスクの低い企業に料率優遇を行うことで、産業界全体の安全性を底上げすることが期待される。これは保険業法 1 条が予定する保険の公共性とも整合する。EIOPA 調査によれば、EU 保険市場において専用 AI ポリシーを策定している保険会社は 49% (2023 年の 25% から 2 倍) に達している。

保険による規律付けの実務上の限界

保険法 28 条は、保険媒介者による告知妨害および不告知教唆があった場合、保険者は告知義務違反を理由とする契約解除ができないと定める。

- **具体的リスク:** 技術的理解の浅い代理店やブローカーが「この程度の AI 利用なら告知書に書かなくても大丈夫」「シャドーAI まで申告する必要はない」といった安易な助言を行えば、それは不告知教唆に該当する可能性がある。こうした場合、保険者は告知義務違反を理由とする解除ができず、規律付け機能は機能しえない。

AI 保険による規律付けの実効性を担保するには、**保険会社自身による直接的なデューデリジェンス**、または**販売チャネルに対する AI リテラシー教育**が必要となる。

Agentic AI と法的責任の帰属

Agentic AI の特性

従来の生成 AI が情報の生成や要約を主たる機能としていたのに対し、Agentic AI は自律的に推論し、計画を立て、システムを操作してタスクを完遂する能力を持つ。エア・カナダ判決が示唆的である。

責任帰属の法的構成

民法上の原則に照らせば、AI を業務に利用している以上、その行為の効果は原則として企業に帰属すると解される。履行補助者としての過失や使用者責任の類推等が考えうる。

しかし、AI の自律性が高まり人間の具体的な指図を超えて行動した場合、その予見可能性の範囲をどう認定するかは難題であり、内部統制システム構築義務と関連する。

むすびにかえて—本稿の結論

1 約款の精緻化

サイバー保険の教訓 (Merck 判決・G&G Oil 判決) を踏まえ、AI 起因の意味を明確に定義し、免責範囲を具体化する必要がある。

2 AI ガバナンス体制構築義務の確立

内部統制システム構築義務の一内容として、組織的ガバナンス枠組みと生成 AI 固有の推論段階における技術的統制 (入力統制・出力監視・権限設計・ログ管理) の四点等が考えられる。

3 保険による規律付けの好循環

AI 保険は AI ガバナンス体制の構築を促す規律付けの機能を担い得る。サイバー保険が情報セキュリティ管理の産業界への浸透を促してきたように、AI 保険も AI ガバナンスの普及を促す制度的な基盤となる。

今後の課題

1

パラメータ設計の実証的検証

AI 保険に特有のモデル精度・システムダウンタイム等のパラメータ設計と損害との相関関係の実証的検証は今後の課題である。

2

AI 除外条項の裁判例の蓄積

海外におけるAI 除外条項の解釈に関する裁判例の蓄積を待って、より精緻な分析を行う必要がある。

3

保険による規律付けの実証的検証

保険による規律付けの実効性について、実証的な検証も求められる。

4

ブラックボックス化する因果関係の法的構成

生成 AI の社会実装が進む中、約款における AI 定義の精緻化、保険の社会的規律付け機能についてさらなる研究が求められている。

本稿が今後、産学の連携をより一層深めた AI 保険の法的基盤の形成に向けた議論の出発点となることを期待する。