AIQMI

AI品質マネジメントイニシアティブ
WG1 SIG-Incident

2025年度AIインシデント事例集

2026年03月31日

目次 / Agenda

はじめに

- 近年AIの利活用が進んでいますが、特に生成AIの利活用は急速に拡大しており、国内外でさまざまなガイドラインが作成されています。
- しかし、これらのガイドラインは抽象度が高く、実際に何をすればよいのか、何を懸念しているのかが分かりにくい場合も少なくありません。
- そこで、実際の事例からガイドラインをみることで、実務におけるガイドライン活用のポイントや検討すべきリスクの参考例をご紹介します。
- 今回は例として、AI事業者ガイドラインを対象として検討を行いました。
- AI品質マネジメントイニシアティブでは、WG1においてさまざまなガイドラインや規制動向、AIインシデントの事例収集などを行っています。

本レポートの位置づけおよび免責事項

はじめに

- 「AIインシデント」については、ガイドラインなどにより表現や定義が異なっています。ここでは、OECDの定義をもとにしています。
- AIインシデント：1つまたは複数のAIシステムの開発、使用、または誤動作が、直接的または間接的に、次のいずれかの被害を引き起こす事象、状況、または一連の事象を指します。
    - 個人または集団への傷害または健康への被害
    - 重要インフラの管理及び運用の混乱
    - 人権侵害、または基本的権利、労働権、知的財産権を保護することを目的とした適用法に基づく義務の違反
    - 財産、コミュニティ、または環境への被害
- 「使用」には、AIシステムの本来の目的外での使用、および意図的または非意図的な誤用から生じる被害も含まれます。

出典：1.Defining AI incidents and related terms, OECD Publishing, 2024,oecd.org/en/publications/defining-ai-incidents-and-related-terms_d1a8d965-en.html（2026年1月5日アクセス）日本語は参考訳

事例1 A社チャットボットの誤案内（AI提供者）

AIチャットボットが生成した誤情報に基づき消費者が損害を被った際、企業が法的責任を負うかどうかが争われた

インシデント情報

1	• 2022年11月、B氏は祖母の死去に伴う航空券購入のため、A社の公式サイトにあるチャットボットを利用した。 • 遺族旅行料金について質問し、チャットボットは割引は遡及（そきゅう）的に適用できると回答した。
2	• その後、B氏は遺族旅行料金についての申請を提出し、A社の担当者から遡及的に適用はできないことを知らされた。 • チャットボットの回答にはA社のウェブページへのリンクが表示されていたが、そのページには遡及適用はできない旨が記載されており、チャットボットの回答とウェブページの内容に矛盾が生じた。
3	• B氏は、全額を支払う必要があることを事前に認識していたら、飛行機は利用しなかったと損害賠償を請求した。 • 裁判においてA社は、チャットボットが提供した情報には責任を負わず、正確な情報は自社ウェブページで入手可能であったと主張した。
4	• 裁判所（Civil Resolution Tribunal）は、B氏は損害賠償を受ける権利があると判断し、A社に損害賠償を支払うように命じた。

※法律的なアドバイスを提供するものではありません。本件はカナダでの事例であり、日本における法令を考慮したものではありません。

出典：2.“Moffatt v. Air Canada, 2024 BCCRT 149 (CanLII),” CanLII, canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html （2026年1月5日アクセス）を基に作成

事例1 A社チャットボットの誤案内（AI提供者）

当事者の主張

- 2022年11月、祖母の死後、 B氏はA社の航空券を予約した。航空券を検索していたB氏は、A社のウェブサイトでチャットボットを利用した。チャットボットは、 B氏にbereavement fares（ビリーブメント運賃）を遡及的に申請できると提案した。 B氏は後にA社の従業員から、A社では遡及的な申請は認められていないことを知らされた。[2]
- A社はB氏が適切な手続きを踏んでビリーブメント運賃を申請しなかったため、遡及して請求することはできないと主張した。また、チャットボットが提供した情報については責任を負わないとした。（Company A says Mr. B did not follow the proper procedure to request bereavement fares and cannot claim them retroactively. Company A says it cannot be held liable for the information provided by the chatbot.）[2]
- B氏によると、A社のチャットボットにビリーブメント運賃について問い合わせたという。チャットボットの回答のスクリーンショットが添付されており、その一部は次のようになっている。[2]
- 「bereavement fares」という語句が、下線付きのハイパーリンクとしてハイライト表示されており、そのリンク先は「Bereavement travel」というタイトルのA社の別ページで、同社のbereavement 規定に関する追加情報が記載されていることには争いがない。（It is undisputed the words “bereavement fares” were a highlighted and underlined hyperlink to a separate Company A webpage titled “Bereavement travel” with additional information about Company A’s bereavement policy.）[2]

証拠に対する分析

- 本件において、サービス提供者と消費者という商業関係に鑑み、A社はB氏に対して注意義務を負っていたと判断する。一般的に、適用される注意義務の基準では、企業は、自社の表明が正確かつ誤解を招かないよう、合理的な注意を払うことが求められている。（Here, given their commercial relationship as a service provider and consumer, I find Company A owed Mr. B a duty of care. Generally, the applicable standard of care requires a company to take reasonable care to ensure their representations are accurate and not misleading.）[2]
- チャットボットにはインタラクティブな要素があるとはいえ、A社のウェブサイトの一部に過ぎない。同社は自社ウェブサイト上の全情報について責任を負うべきであり、その情報が静的なページから提供されるかチャットボットから提供されるかは問題ではない。（While a chatbot has an interactive component, it is still just a part of Company A’s website. It should be obvious to Company A that it is responsible for all the information on its website. It makes no difference whether the information comes from a static page or a chatbot.）[2]
- A社は、 B氏がウェブサイトの別の場所で正しい情報を見つけることができたと主張しているが、「遺族旅行」というタイトルのウェブページがチャットボットよりも本質的に信頼できる理由を説明していない。また、顧客がウェブサイトのある部分で見つけた情報を、別の部分で再確認しなければならない理由も説明されていない。（While Company A argues Mr. B could find the correct information on another part of its website, it does not explain why the webpage titled “Bereavement travel” was inherently more trustworthy than its chatbot. It also does not explain why customers should have to double-check information found in one part of its website on another part of its website.）[2]

※法律的なアドバイスを提供するものではありません。本件はカナダでの事例であり、日本における法令を考慮したものではありません。

出典：2.“Moffatt v. Air Canada, 2024 BCCRT 149 (CanLII),” CanLII, canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html （2026年1月5日アクセス）日本語は参考訳、A社およびB氏の伏せ字は筆者加工

事例1 A社チャットボットの誤案内（AI提供者）

体制面での検討

- 生成AIは確率的に動作し、当初の想定と異なる挙動が生じる可能性があるため、インシデント発生時の対応を事前に検討しておくことが望ましい。

個別システムの検討

- 本事例では、 webページに掲載されていたポリシーとAIチャットボットが回答した内容が矛盾していた。
- チャットボット運用後もモデルの変更や、RAGにおいては利用データの変更もあることから、定期的なモニタリングが望ましい。

AI事業者ガイドラインでの記載（代表的な記載の項目名のみを抜粋しています。網羅的ではありません）

- 行動目標 3-1-2【AI 利用者及び業務外利用者に対する、乖離可能性/対応策に関する十分な情報提供】[4]
- 行動目標 3-4【予防・早期対応による AI 利用者及び業務外利用者のインシデント関連の負担軽減】[4]
- 行動目標 3-4-1【各主体間の不確実性への対応負担の分配】[4]
- 行動目標 3-4-2【インシデント発生時の対応の事前検討】[4]
- 外部のAI利用者に対して問い合わせ先やAIが用いられていることを明記し、内部では責任者や連絡体制を整備することが望ましい。
- 行動目標 3-3-2【環境・リスク分析のための日常的な情報収集・意見交換の奨励】[4]
- AIの利用は発展途上であり、リスクの種別や社会の受容性が変化することが想定される。日常的にガイドラインやインシデント等の情報を収集し、社内で共有・検討をすることが望ましい。[3]

- 行動目標 4-2【個々の AI システム運用状況の説明可能な状態の確保】[4]
- AIシステムは、運用中に挙動が変化する要素が多いため、AIの運用状況を継続的にモニタリングすることが望ましい。
- 参考：行政の進化と革新のための生成 AI の調達・利活用に係るガイドライン
    - 利用ログが取得できる場合の例として、プロンプトや出力結果をサンプルチェックし、定期的にモニタリングすることが示されている。[5]

出典：3.総務省・経済産業省「AI 事業者ガイドライン（第1.1版）」、meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_1.pdf （2026年1月5日アクセス）
4.総務省・経済産業省「AI 事業者ガイドライン（第1.1版）別添（付属資料）」、meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_3.pdf （2026年1月5日アクセス）
5.デジタル庁「行政の進化と革新のための生成 AI の調達・利活用に係るガイドライン」、digital.go.jp/assets/contents/node/basic_page/field_ref_resources/e2a06143-ed29-4f1d-9c31-0f06fca67afc/80419aea/20250527_resources_standard_guidelines_guideline_01.pdf（2026年1月5日アクセス）

事例2 C社チャットボットの想定外利用（AI提供者）

C社のAIチャットボットが、本来提供するカスタマーサービス外であるプログラミングコードの生成などに利用され、想定外利用や攻撃試行が行われた

インシデント情報

1
- C社では、遺失物の問い合わせなどをはじめ、AIを用いたチャットボットによるカスタマーサービスを提供していた。
- 一部ユーザーがチャットボットにプログラムの作成を指示したところ、本来のカスタマーサービス外であるプログラミングコード（挿入ソートの一部コード）が提示された。なお、初期に投稿された画像では、Jailbreakなどガードレールを無効化するような指示は含まれず、プログラミングの記述のみが指示されていた。

2
- その後、ユーザーがSNSに結果を画像で投稿したところ、その投稿を見た人々がチャットボットに対してさまざまな依頼を指示し、投稿した。

3
- さらに一部のユーザーでは、応答規則の抽出や制約回避を試みる投稿もみられた。

出典：6.Mia, “權限沒設好？網友測試北捷 AI 客服還能寫程式，官方：勿濫用公共資源,” INSIDE, 24 November 2024, inside.com.tw/article/36868-taipei-mrt-ai-customer-service（2026年1月5日アクセス）を基に作成

事例2 C社チャットボットの想定外利用（AI提供者）

体制面での検討

- 事例1にて記載の内容を含め、セキュリティインシデントが発生した場合の対応も検討しておくことが望ましい。

個別システムの検討

- 生成AIの機能を切り離しても運用できるよう検討しておくことが望ましい。
- 異常な利用発生時に被害を最小限とするため、トランザクション量の監視を行うことが望ましい。

AI事業者ガイドラインでの記載（代表的な記載の項目名のみを抜粋しています。網羅的ではありません）

- 事例1と類似の取り組みは省略
- P-2）ii. 適正利用に資する提供（第4部 AI 提供者に関する事項 AI システム・サービス提供後）[3]
    - セキュリティインシデントが発生した場合の対策を検討しておくことが望ましい。

- P-5）i. セキュリティ対策のための仕組みの導入（第4部 AI 提供者に関する事項 AI システム実装時）[3]
- P-5）ii. 脆弱性への対応（第4部 AI 提供者に関する事項 AI システム・サービス提供後）[3]
    - AIシステムが適正に利用されているかモニタリングするとともに、システムに異常が発生した場合のインシデント対応やAIシステムの切り離し・停止を考慮することが望ましいと考えられる。
- AI ライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度な AI システムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じる（第2部 AI により目指すべき社会及び各主体が取り組む事項 D. 高度な AI システムに関係する事業者に共通の指針）[3]
    - 高度なAIシステムを扱うAI提供者については、適切な範囲でレッドチームなどの手法を組み合わせたテストを実施することが望ましいとされている。

出典：3.総務省・経済産業省「AI 事業者ガイドライン（第1.1版）」、meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_1.pdf （2026年1月5日アクセス）

事例3 D社の存在しない書籍を含む読書リスト掲載（AI利用者）

D社が発行した特集記事について、おすすめ書籍リスト15冊のうち10冊のタイトルと説明が虚偽または全くの架空であった

インシデント情報

1
- D社は定期購読者向けに特別版の発行をしており、夏のおすすめ書籍リストを掲載した。
- 記事はコンテンツパートナーからライセンス供与を受けたもので、コンテンツパートナーはフリーランスのコンテンツクリエーターと協力して記事を作成していた。
- コンテンツパートナーによれば、フリーランスのコンテンツクリエーターはコンテンツ作成ポリシーに違反し、AIエージェントを利用して記事を作成した。

2
- その結果、15冊の書籍のうち10冊のタイトルと説明が虚偽または全くの架空のものであった。
- コンテンツクリエーターは記事をチェックせずに記事を送付し、コンテンツパートナーも十分な事実確認や編集を行わないままD社に送付した。

3
- D社においても記事の十分な事実確認や編集が行われないまま掲載され、コンテンツが第三者によって作成されたことも明記されていなかった。

4
- その後の調査により、実名の記載がある他のセクションの記事にも、誤りや独自に検証できない誤情報が含まれていることが発覚した。

出典：7.Melissa Bell, “Lessons (and an apology) from the Sun-Times CEO on that AI-generated book list,” CHICAGO SUN-TIMES, 30 May 2025, chicago.suntimes.com/opinion/2025/05/29/lessons-apology-from-sun-times-ceo-ai-generated-book-list（2026年1月5日アクセス）を基に作成
8.Dan Mihalopoulos, “Special section with fake book list plagued with additional errors, Sun-Times review finds,” CHICAGO SUN-TIMES, 30 May 2025, chicago.suntimes.com/news/2025/05/29/special-section-king-fake-book-list-errors-sun-times-review（2026年1月5日アクセス）を基に作成

事例3 D社の存在しない書籍を含む読書リスト掲載（AI利用者）

体制面での検討

- 自社のAIシステムだけではなく、サプライチェーン全体でAIを利用した出力結果が利用されることを考慮することが望ましい。
- 既存と同様の業務やコンテンツについても、AIの利用が想定されるため、改めてリスク評価や対策の検討を行うことが望ましい。

AI事業者ガイドラインでの記載（代表的な記載の項目名のみを抜粋しています。網羅的ではありません）

- U-2）i. 安全を考慮した適正利用（第5部 AI 利用者に関する事項 AI システム・サービス利用時）[3]
  - AIの出力についてリスクを認識した上での利用が記載されている。
- U-7）i.関連するステークホルダーへの説明（第5部 AI 利用者に関する事項 AI システム・サービス利用時）[3]
  - AIを利用した出力結果がサプライチェーンのいずれかの部分で利用される可能性がある。ステークホルダー間で取り扱いについて合意することが望ましい。
- 行動目標 1-1【便益/リスクの理解】（A.経営層による AI ガバナンスの構築及びモニタリング 1.環境・リスク分析）[4]
- 行動目標 3-3-2【環境・リスク分析のための日常的な情報収集・意見交換の奨励】[4]
- 行動目標 6-1【行動目標 1-1～1-3 の適時の再実施】[4]
  - ハルシネーションや著作権など、AIによって顕在化するリスクの最新動向を把握し、自社の取り組みを検討することが望ましい。特に、本事例のように、既存の業務プロセスにAIエージェントなどの利用が導入される場合、新規のAIシステムの導入とは異なる観点で評価が必要となる可能性がある。

出典：3.総務省・経済産業省「AI 事業者ガイドライン（第1.1版）」、meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_1.pdf （2026年1月5日アクセス）
4.総務省・経済産業省「AI 事業者ガイドライン（第1.1版）別添（付属資料）」、meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20250328_3.pdf （2026年1月5日アクセス）

AI品質マネジメントイニシアティブ