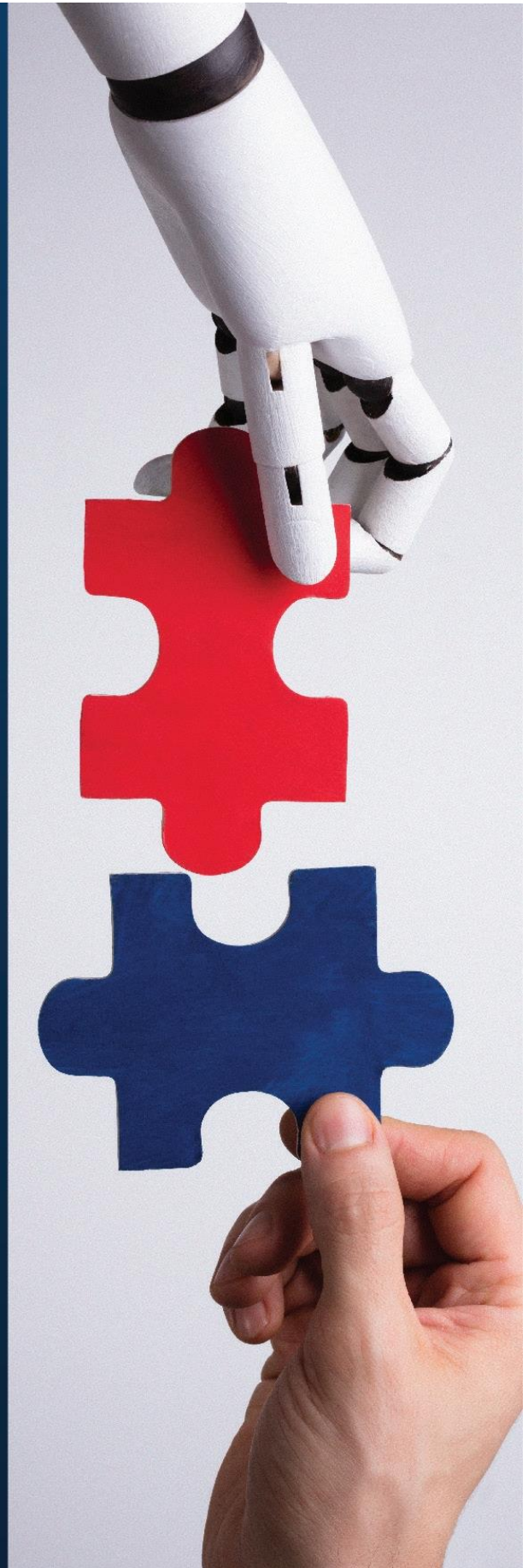


市場調査・インサイトのための
AI 活用サービス購入者に役立つ
20 の質問

ESOMAR

20 Questions
to Help Buyers of
AI-Based Services
for Market Research
and Insights



ESOMAR について

1947 年の設立以来、ESOMAR は調査・インサイト・および分析のグローバルハブとなっている。5 万人以上の個人、750 社以上の法人、130 カ国以上に広がる世界規模の会員組織であり、インサイトの専門家と企業が、グローバルとローカルの両方の舞台で潜在能力を発揮できるように支援し、つながり、協力し、成長し、知識を共有することを促進している。

包括性、思いやり、イノベーション、信頼というコアバリューに基づき、75 年以上にわたって急速に進化する業界をリードしてきた。専門的・倫理的規範の向上、教育の促進、立法者への提言、ベストプラクティスの共有、エビデンスに基づく意思決定者のためのソリューションの推進に引き続き取り組んでいる。www.esomar.org

目次

本ガイダンスの目的と範囲	1
A. 会社概要	2
B. AI の能力／サービスは 説明可能で目的にかなっているか？	3
C. AI の能力／サービスは 信頼でき、倫理性、透明性があるか？	4
D. 人間による AI システムの監視を どのように行うか？	5
E. データガバナンス・プロトコルは どのようにになっているか？	7
用語集	10
プロジェクトチーム	12

本ガイダンスの目的と範囲

このチェックリストの目的は、AI を活用したサービスを依頼しようとしている市場・世論・社会調査のリサーチャーとデータアナリストのための、**ガイダンス**として役立つことである。すなわち、AI サービスの購入者とサプライヤー（供給者）が、この分野における適切なレベルの透明性、信頼、信用を確保し、適用される知的財産法、プライバシー法、AI 法が尊重されるようにすることである。その意図は、市場・世論・社会調査およびデータ分析（以下、「調査」と呼ぶ）において公正で、透明性、責任性があり、倫理的な AI の利用を促進することである。

調査における現在の AI 活動の範囲に関して、私たちには考慮すべき 3 つの分野がある。

1. 一般化されたビジネスプロセスなど、調査分野の外部で開発された、ビジネスそのものに関連する AI ツールの使用。
2. 調査活動を実施する際の AI の使用。例えば、サプライヤーによって導入された商業的な調査や、購入者向けの特注品など。これには、特定の調査製品やサービスの開発、調査組織が調査の参加者とやり取りする方法、調査活動で使用される方法、調査の結論の準備と事後分析、独自のシンジケートデータセットの開発と分析に使用される方法などが含まれる。
3. 第三者の AI プログラムにおける調査データおよび調査結果の使用。これには、AI の「トレーニングデータ」としての調査データの使用を含む。これらの第三者とは通常、調査プログラムのクライアントや委託者である。

このガイダンスの範囲は 2 番目の分野をカバーし、1 番目と 3 番目は除外している。

この急速に発展している分野に倫理的・専門的原則を適用しようとしているのは、調査部門だけではない。そのため、プロジェクトチームは、このガイダンスを作成する際に、様々な既存の倫理規定と、予想される法的展望に関する情報を併せて参照した。

AI の定義自体は、多くの立法当局者がいまだに議論している問題である。この文脈において本資料では、AI ベースの調査ソリューションを提供すると主張するあらゆる組織、または AI が利用されているとユーザーが確信し、このフレームワークが役立つ可能性があるあらゆる状況で使用できるように設計されている。用語に関しては、既存の IAPP¹ の AI 定義用語集を参照した。本資料に関連する IAPP の用語は、完全な定義用語集へのリンクおよび IAPP の著作権義務とともに、附属書で公開されている。

このプロジェクトのワーキンググループは、このトピックのダイナミックな性質を認識している。草案の作成時点では、調査のほぼすべての領域が AI ベースの活動の候補として議論されており、購入者とサプライヤーが急速な変化に遅れをとらないように、その影響を理解しようとしている。多くの新規および既存の組織が AI を活用したソリューションを提供しており、その中には業界を変革する能力を持つものもある。

このガイダンスは、短期間のうちに時代遅れになる可能性があり、AI 分野が驚くべき速さで加速し、多様化しているため、一部の組織にとっては他の組織よりも関連性が高く、特定の AI アプリケーションとの関連性が高い可能性があることが理解されている。それでも、まずはどこかで始める必要がある。これらすべての点を認識し、この資料はダイナミックなビジネス環境と法的環境を考慮し、随時更新される予定である。

¹ IAPP = *International Association of Privacy Professionals*

A. 会社概要

このセクションの内容

このセクションは、サプライヤー組織の信頼性を最初に理解するのに役立つ。

1. 調査にAIベースのソリューションを提供する上で、貴社にはどのような経験とノウハウがあるか？

コンテキスト

この回答は、調査とAIの両方の分野でAIベースのソリューションを提供するサプライヤーの経験と知識について、購入者が認識を形成するのに役立つ。たとえば、そのサプライヤーは調査のスキルと、データサイエンス、ディープラーニング、自然言語処理などのAI関連のスキルを適切に組み合わせ持っているか？この議論は、購入者が提供される専門知識のレベルと深さを理解するのに役立つであろう。

2. AIベースのサービスが調査によい影響を与える可能性があるのはどのような場合だと思うか？AIはどのような特徴と利点をもたらし、どのような問題に対処するのか？

コンテキスト

この質問は、AIベースのソリューションが調査の既知の問題を解決し、調査のプロセスと意思決定や社会への貢献を改善する方法について、サプライヤーがどの程度考えているかを示すものである。提供されているサービスについてより詳細に議論する前に、AIツールの使用に関するサプライヤーの一般的な見解を理解しておく役立つ場合がある。

3. AIの使用と展開において、どのような実際的な問題や課題に直面したか？うまくいったこととその方法、うまくいかなかったこととその理由は何か？

コンテキスト

この回答は、購入者がサプライヤーの組織内の活動レベルと学習内容を評価するのに役立つ。また、新しいテクノロジーに関する課題を管理する上で、サプライヤーがどの程度の経験を積んでいるかを示す指標も提供される。

B. AI の能力／サービスは説明可能で目的にかなっているか？

このセクションの内容

このセクションは、購入者が実用的な観点から AI サービスを評価するのに役立つ。これにより購入者は、提供される能力がビジネス目的に沿っているかどうか、明確なメリットを提供する可能性が高いかどうかを判断できる。

4. サービス提供における AI の役割を、リサーチャーや利害関係者が簡単に理解できるように、専門的ではないシンプルな言葉で説明できるか？ 主要な機能は何か？

コンテキスト

説明のしやすさは、方法論が広い意味でどのように機能するのか、方法論を強化するために AI がどのように使用されるのか、そのアプローチがもたらす利益を理解するための鍵である。

5. 使用されている AI モデルは何か？ 御社の AI ソリューションは、主に社内で開発されているか？ 既存の AI システムを統合しているか？ サードパーティが関与しているか？ そうであれば、それらはどのように？

コンテキスト

購入者は、能力と潜在的なリスクを評価できるように、カスタムビルドとオープンソースソリューションを区別することが役立つ場合がある。一般に入手可能なソリューションが、サプライヤーの製品に統合されているかどうかを購入者に明確にするべきである。これにより、機密情報の取り扱いなど、潜在的な影響を評価できる。

6. 導入されたアルゴリズムは、どのようにして望ましい結果を提供するか？ AI サービスをトレーニングするための基礎データと、モデルとの相互作用の方法を要約できるか？

コンテキスト

使用されるデータ、データの使用方法、およびそのデータに適用されるアルゴリズム処理に関する情報は、AI が要求される目的に適合しているかどうかを判断するのに役立つ。

トレーニングデータとしてクライアントのデータや合成データが使用されているかどうか、または必要であるかどうか、および購入者がこれをオプトアウトできるかどうかも理解に役立つ。多くの AI システムは、取り込んだデータから継続的に学習する。これはユーザーにとって有益であるが、データの出所と信憑性を考慮するべきである。

ユースケースや関心のあるサービスに応じて、モデルが英語以外のコンテンツを使用してトレーニングされているかどうかを調べることに役立つ場合がある。

C. AI の能力／サービスは信頼でき、倫理性、透明性があるか？

このセクションの内容

このセクションは、購入者とサプライヤーが倫理原則に沿っているかどうか、サプライヤーが潜在的なバイアス、データセキュリティ、回復力などの他の重要なトピックを考慮しているかどうかを明確にするのに役立つ。

7. アウトプットの正確性を検証し、妥当性を確認するプロセスはどのようなもので、かつ文書化されているか？ 妥当性をどのように測定して評価するか？ システムが信頼性に欠ける、ゆがんだ、または偏った結果をもたらす場合を特定して対処するプロセスはあるか？ アウトプットを微調整するために特定の手法を使用しているか？ 生成された結果が "目的にかなっている" ことをどのように確認するか？

コンテキスト

AI には、目的によって、バイアスや代表性に関連する課題や、一貫性のない結果や不正確な結果を生成するリスクがある。この質問の目的は、信頼できる結果を保証するためにサプライヤーが使用しているプロセスを理解することである。たとえば、回答では、トレーニング用データセットがどのように選択されているか、何をカバーしているか、システムが導入される環境を代表するために十分に最新であるかどうか、スクリーニングされているかどうかを示すべきである。

生成 AI と合成データに関する補足的な質問と、その背景を以下に示す。

- 生成 AI を使用するツール: AI は、現実世界や特定の日付以降のできごとに関する知識が限られている可能性があり、時には偏った、不正確な、または誤解を招くような回答を生成することがある。そのようなツールは AI モデルにどの程度組み込まれており、「ハルシネーション (幻覚)」や不正確な回答を特定し、バイアスを最小化するためにどのような対策が取られているか。
- 合成データを使用するツール: 合成データのアウトプットを、調査の主要なアウトプットまたは実際の結果と比較するために、どのような検証が行われているか。データの最新性に関する基準はあるか。データが使用に適しているかどうかを評価できるか。サプライヤーはそのシステムやアプローチにおいて、自然人から直接得られたデータと合成的に得られたデータとを、どのように区別しているか。

8. その AI モデルにはどのような限界があり、それらをどのように軽減しているか。

コンテキスト

サービスの限界を理解することは、一貫性のない、または不正確なアウトプットの潜在的なリスクを判断するために不可欠である。この質問の目的は、サプライヤーが提案するサービスの限界について透明性があるかどうかを評価することである。購入者には、結果および導き出された結論の妥当性を評価できるように、十分な技術情報が提供されるべきである。リスクの軽減策には、品質評価基準、テスト、承認プロセスに加えて、新たな問題を特定して管理するための、合意された基準に対する定期的なレビューが含まれる場合がある。

9. 人間に対する注意義務を念頭に置いてサービスを設計するために、どのような事項を考慮したか？

コンテキスト

市場調査では、倫理性が重要な役割を果たしている。この質問は、サプライヤーがそのサービス設計の中で、AI モデルを人間に使用することで潜在的な悪影響が生じる可能性があるかどうか、考慮したかを評価するのに役立つ (例えば、偏見の増加、金銭的危害、誤った情報など)。

D. 人間による AI システムの監視をどのように行うか？

このセクションの内容

このセクションは、購入者が、提供されている AI アプリケーションの開発と運用の両方で、人間の関与と監視がどのように考慮されているかを理解するのに役立つ。購入者は、サプライヤーが、そのプロセスにおける人間の監視について、またはその手法のユーザーがアウトプットをストレステストできる方法について説明できることを期待すべきである。

このセクションの質問に対する回答は、AI を活用してソリューションを構築し、倫理的かつ責任ある方法で AI が処理/分析するデータを扱うときに、人間がどのような役割を果たすかを特定する方法を決定するのに役立つ。

10. 透明性: どのようにして、サービスのどの部分で AI 技術が使用されていることを明らかにするのか？

コンテキスト

AI が生成した画像/テキストを特定することは、購入者にアウトプットの原典を示すことに役立つ。サプライヤーは、サービスで使用されている AI 技術の種類を伝えることができるようにすべきである。たとえば、関連する場合、サプライヤーは画像/テキストが「AI によって生成された」ことを示すフラグを付けることができる。

11. AI を活用したソリューションに対して明確に定義された倫理原則があるか？ また、実際に AI の動作を決定するのにどのように役立つか？ 人間が定義した倫理原則が、AI 主導型ソリューションを背後から支配していることをどのように保証するか？

コンテキスト

この質問は、サプライヤーが倫理原則を定めているかどうかを確認し、ソフトウェア・モジュールをテストするときに人間がどの程度関与しているかを調べるように設計されている。

たとえば、人間によるフィードバックを使用したモデルの強化学習によって AI システムをトレーニングすることは、AI モデルを使用または調整するための主な手法であり、バイアスの発生や強化を回避するのに役立つ。購入者は、AI によって生成されたアウトプットが意思決定に使用される場合や、個人に影響を与える可能性がある場合に、人間による検証とチェックが行われていることを知る必要があるだろう。

12. 責任あるイノベーション: 倫理的なコンプライアンスを遵守するために、AI ソリューションに人間の監視をどのように組み込んでいるか？

コンテキスト

この質問により、購入者は、サプライヤーが調査向け AI ソリューションの設計と開発において、人間が関与するツール、手法、プロセスを組み込んでいるかどうかを理解することができる。これらの介入の程度は、AI サプライヤー組織の規模や性質、提供されるサービスまたは AI の種類によって異なる。例えば、以下のようなものがある。

- ヒューマン・イン・ザ・ループ²: このアプローチでは、機械学習アルゴリズムの構築と更新のプロセス全体に人間が関与することを求めている。

² ヒューマン・イン・ザ・ループ (HITL) - AI ベースのシステムを利用する文脈では、HITL はプロセスに人間の判断を統合するためのフレームワークを指す。これには、継続的な監視、人間による無効化、人間による最終決定、例外処理などが含まれるが、これらに限定されない。出典: ESOMAR。

- 倫理審査委員会: 多様なバックグラウンドを持つ専門家で構成される、独立した倫理審査委員会を設置する。
- 参加型設計プロセス: 欧米以外のコミュニティを含む多様な利害関係者を AI ソリューションの設計プロセスに関与させ、ソリューションがさまざまなコミュニティのニーズに適合されるようにする。
- 文化的感受性トレーニング: AI ソリューションに取り組む開発者やデータサイエンティストに文化的感受性トレーニングを提供し、文化の違いを理解し、より包括的なソリューションを開発できるようにする。
- 人間によるデータキュレーション: ナレッジグラフと呼ばれる詳細なネットワークの作成が含まれる。これらのグラフの各点はデータを保持し、すべての接続は意味のある関係を表している。ユーザーは、厳選されたデータを使用して、これらのグラフに慎重に選択した情報を追加することができる。これにより、大規模言語モデル (LLM) 駆動クエリを使用した貴重なコンテキストが提供される。機械によって作成された追加データは、人間の専門家によってチェックされるべきである。
- 人間が設計した仕様 (Human-engineered ontology) : これらの組織化されたモデルを使用して、ユーザーが機械の考え方に影響を与えることで、LLM を導き、その行動に制限を設けることに役立つ。これらのモデルに変更を加える場合には、ソリューションとして使用する前に人間によるチェックがなされる必要がある。

E. データガバナンス・プロトコルはどのようになっているか？

このセクションは、このガイダンスの他の質問と組み合わせて使用することも、独立したチェックリストとして使用することもできる。そのため、この資料の他のセクションと重複する要素があるかもしれない。

このセクションの内容

このセクションは、サプライヤーが AI ベースの活動を管理する法的枠組みを適切に認識しているかどうかを、バイヤーが理解するのに役立つ。

AI サプライヤーとそのクライアントは、データ保護法や規制によって課されるデータ保護および関連する情報セキュリティ要件の対象となる。これらの法規制は、行政区域（国や、同一国内でも州）によって適用される法律や規制が異なる。また、一般的には、データが収集された場所またはプロバイダーの所在地に基づいて解釈される³。

さらに、知的財産権と著作権に関連する法律や規制の対象となる場合もある。

このセクションの質問に対する回答は、購入者が、サプライヤーが実装したデータ保護、情報セキュリティ、コンプライアンスのポリシー、手順、および慣行を理解するのに役立つ。

13. データの品質: AI モデルに使用されるトレーニングデータが正確で完全であり、信頼できる結果のために調査の目標に関連しているかどうか、また、データプライバシー法で要求されているかどうかをどのように評価するか？

コンテキスト

不正確、不完全、または無関係なデータは、AI からの結果に悪影響を与える可能性がある。AI のデータ品質は、AI が期待どおりに動作することを保証するための重要なステップである。データの品質が低いと、アウトプット、インサイト、および結論が妥当でなくなる可能性がある。たとえば、LLM の中でさえ、特定の国籍または国籍内の人口統計は過小評価される可能性がある。より広い人口統計に関連するインサイトを決定する際に、これはどのように考慮されているのか。

14. データの系統: トレーニングデータまたは入力データの出所と処理方法を文書化しているか？ また、これらの情報源は（今も）利用可能か？

コンテキスト

データはどこから入手したのか、誰がどのように処理したのか、独自のデータなのか一般に公開されているデータなのかを理解することが重要である。購入者は、データが責任を持って倫理的に調達・処理され、AI システムの目的に適合していることを理解したいと考えるであろう。

³ 適用されるデータ保護法や規制には、以下のものが含まれるがこれらに限定されない。個人情報保護に関する法律(日本); オーストラリアプライバシー法 (オーストラリア); カリフォルニア州消費者保護法 (米国カリフォルニア州); 児童オンラインプライバシー保護法 COPPA (米国); 個人情報保護法 PIPL (中国); データ保護法 (英国); データ保護法のデータローカライズ要件に関する改正 (ロシア連邦); 一般データ保護法 (ブラジル); EU 一般データ保護規則または EU-GDPR (EU/EEA); 医療保険の相互運用性と説明責任に関する法律 HIPAA (米国); Graham-Leach Bliley 法または GLBA (米国); および PIPEDA (カナダ)。

15. プライバシー通知 (プライバシーポリシーとも呼ぶ) へのリンクを提供してください。貴社が製品またはサービスごとに異なるプライバシー通知を使用している場合には、この質問への回答で対象となる製品またはサービスに関連する例を示してください。

コンテキスト

プライバシー通知は、データ保護に関する法律や規制、その他の法律や規制、および多くの市場調査業界の規範によって義務付けられている。サプライヤーが収集・処理する個人データ、個人データの使用目的、個人データの使用・開示・管理方法に関する情報を開示するべきである。サプライヤーのプライバシー通知を確認することで、購入者は、個人データに関連する手順と慣行、および適用される法律・規制・業界規範をどの程度遵守しているかを理解することができる。

16. データ保護法を遵守し、調査参加者のプライバシーを保護するための対策を実施するために、どのような手順を踏んでいるか？ プライバシー法で要求されている個人に対するリスクを評価し、必要な場合に、または別の法的根拠によって、データ処理のために確実に同意を取得しているか？

コンテキスト

サプライヤーのコンプライアンス上の立場を理解することは、法律を確実に遵守し、個人データを保護するために不可欠である。

世界的なプライバシー法では、リサーチャーはAIが人々にもたらす可能性のあるリスクを評価し、多くの場合、特定の目的でデータを処理するために同意を取得することが求められている。これは通常、データ保護影響評価 (DPIA) を通じて行われ、各国のデータ保護規制当局からガイダンスが提供されている。

人々に関連する個人データが処理される場合、プライバシー法では通常、リサーチャーが個人に対して、データの処理方法、共有相手、使用目的、AIで使用される際の安全性の保持方法について、透明性を確保することを求めている。購入者とサプライヤーは、データ保護法と規制によって課される、データ保護と関連する情報セキュリティ要件の対象となる。

17. 敵対的な攻撃、ノイズ、その他の潜在的な混乱に対するAIシステムの回復力 (レジリエンス) を確保するために、どのような手順を踏んでいるか？ また、どのような情報セキュリティのフレームワークと基準⁴を使用しているか？

コンテキスト

回答には、AIシステム、その設計、またはトレーニングデータの操作などの技術的なミス、またはシステムの誤用や不適切な使用の可能性など、脆弱性と潜在的な脅威を特定して、対処するプロセスを含めるべきである。また、サイバー攻撃への対抗、サイバーセキュリティシステムおよびモデルが特定のセキュリティ基準に準拠しているかどうかを検出して対処するための対策も含めるべきである。問題が発生した場合のバックアップ計画を可能にする、緊急時の代替プランはあるか。データがどこに存在するのか、クラウドとデバイスのどちらに保存されているかを理解することが役立つ場合がある。

⁴ 例としては、COBIT、HITRUST、ISO 27001、NIST Cybersecurity Framework、SOC 2 などがあるが、これらに限定されない。

18. データの所有権: 知的財産権や使用許諾などを含む、データの所有権を明確に定義し、伝えているか?

コンテキスト

モデルに入力されているデータの所有者は明確か? 入力されるデータは、自分で作成した情報である場合もあれば、クライアントや別のサプライヤーから入手した情報であるかもしれない。サプライヤーまたはクライアントがデータを所有している場合は、データの所有権を保持し、データの使用方法を管理している可能性がある (たとえば、ChatGPT やその他の同等のモデルにアップロードしない)。最初に、データで実行できることとできないことを明確にし、それを適切に文書化することをお勧めする。

データ所有者は、情報を機密または秘密にしておきたいと望み、パブリックドメインや AI システムのクラウドベースのツールに公開しないように要求するかもしれない。どのデータが機密と見なされ、公開されているテクノロジーにアップロードされるべきでないかを明確にすることが重要である。

19. データ主権: データを使ってできることを制限しているか?

コンテキスト

データ所有者は、自分のデータが AI によって処理されることを快く思っていないかもしれない。多くの AI テクノロジーは新しいものであり、データ所有者は、自分が認識していない、確認していない、または信頼していないシステムに自分のデータをアップロードしたくない場合がある。データ所有者は、自分のデータが他のデータセットと混在することを許可しない場合がある。さまざまな新しい AI 法があり、データ所有者は、準拠していない可能性のある AI テクノロジーに自分のデータが読み込まれることを望まないであろう。データ学習や AI ツールを使用したアウトプットの生成において、クライアントが自分のデータを AI で処理することを許可するか禁止するかを契約書に含めることをお勧めする。データ所有者は、データやモデルを保持すべき国や地域を規定する必要があるかもしれない。

データ所有者はまた、AI モデルへの入力 (プロンプト) に含まれる知的財産を所有している当事者を確認すべきである。データ所有者は、データやモデルを保持すべき国や地域を規定する必要があるかもしれない。

20. (アウトプットの) 所有権: 誰がアウトプットの所有者になるかを明確にしているか?

コンテキスト

データ所有者は、どの当事者がアウトプットの権利を所有することになるのか、つまり、AI テクノロジーによって生成された情報の知的財産 (IP) を誰が所有するかについて懸念している可能性がある。たとえば、サードパーティの AI テクノロジー所有者は、AI テクノロジーからのアウトプットの IP を所有することを利用規約に定めている場合がある。このことによって、特に意図されている利用が商業的な性質のものである場合に、情報の使用方法が制限される可能性がある。

用語集

本資料の作成に使用された用語集は、IAPP (国際プライバシー専門家協会 : International Association of Privacy Professionals) から提供されている。IAPP が作成した “AI ガバナンスのための主要用語集” は、IAPP リソースセンターに掲載されている。許可を得てここに転載している。このバージョンは 2023 年 6 月に発行された。

以下の用語は、この ESOMAR ガイダンスに関連する用語集のサブセットである。

Artificial Intelligence; 人工知能⁵

人工知能とは、さまざまな計算技術を使用してタスクを実行または自動化する、工学的なシステムを説明するために使用される広義の用語である。これには、機械が経験から学習し、新しい入力データに適応し、以前に人間が行ったタスクを実行する可能性がある、機械学習などの技術が含まれることがある。より具体的には、コンピューターの知的な動作をシミュレートすることに特化したコンピューターサイエンスの一分野である。意思決定の自動化が含まれる場合もある。略語 (頭字語、以下同様) : AI

Bias; バイアス

AI 分野にはいくつかの種類バイアスがある。計算バイアスとは、モデルの仮定や入力データ自体に起因する、予測の真の値からの系統誤差または偏差である。認知バイアスは、不正確な個人の判断や歪んだ思考を指し、社会的バイアスは、個人または集団に対する組織的な偏見、好意、または差別をもたらす。バイアスは結果に影響を与え、個人の権利と自由にリスクをもたらす可能性がある。

【日本語版注記】

AI 分野で一般に使用されている「Bias」には、さらに別の種類がある。ディープラーニングでの慣用語として使用される場合、ニューラルネットワークの要素である Bias は、ユニット固有の「閾値」という意味でモデルに現れ、その閾値を超えて初めて信号が次のユニットに伝送される。その Bias と同じ効果を回帰分析では定数項と呼んでいる。

Explainability; 説明可能性

AI システムが特定のアウトプットを生成する方法、または特定の文脈で判断に到達する方法について、事前に定められた相手先に対して十分な情報を記述または提供する能力。XAI は、AI の透明性と信頼を維持するために重要である。略語: XAI

Generative AI; 生成 AI

大規模なデータセットで学習させたディープラーニングを使用して、記述されたテキスト、プログラムコード、画像、音楽、シミュレーション、動画などの新しいコンテンツを作成する AI の分野。判別モデルとは異なり、生成 AI は新しいデータではなく、既存のデータに基づいて予測を行う。これらのモデルは、入力データやユーザーのプロンプトに基づいて新しいアウトプットを生成することができる。

Hallucinations; ハルシネーション (幻覚)

生成 AI モデルが、情報ソースと矛盾するコンテンツを作成したり、あたかも事実であるかのように見えるが、事実とは異なるアウトプットを作成したりする事例。

⁵ 経済協力開発機構 (OECD) は、AI を次のように定義している。「AI システムとは、明示的または暗黙的な目的のために、物理的または仮想的な環境に影響を与える可能性のある予測、コンテンツ、推奨、決定などのアウトプットを生成する方法を、受け取った入力から推論する機械ベースのシステムである。運用開始後の自律性と適応性のレベルは、AI システムによって異なる。EU の AI 法における AI の定義は同じであるが、最後の一文が除外されている。

Input data; 入力データ

アウトプットを生成する目的で、学習アルゴリズムまたは機械学習モデルに提供されたデータ、またはそれらのモデルによって直接取得されたデータ。機械学習モデルが学習、予測、またはタスクを実行するための基礎となる。

Large Language Model; LLM, 大規模言語モデル

深層学習アルゴリズムを利用して、大量のテキストデータセットで事前学習させたモデル（「機械学習」を参照）を作成する AI の一形態。言語学習の一般的な目的で、文字、単語、フレーズ間のパターンと関係性を分析・学習する。LLM には一般的に 2 つのタイプがある。学習データから学んだ単語列の確率に基づいて次のテキストを予測する生成モデル（「生成 AI」を参照）と、データの特徴や重みの確率に基づいて分類を予測する判別モデルである。「大規模」という用語は、一般的に、パラメーターの数によって測定されるモデルの容量と、トレーニングに使用される膨大なデータセットを指す。略語: LLM

Machine Learning; 機械学習

コンピューターシステムが入力データから反復的に学習し、入力データに基づいて判断、推論、または予測を行うことを可能にするアルゴリズムを含む、AI のサブ分野である。これらのアルゴリズムは、新しいデータに対して特定のタスクを実行するように明示的にプログラムされていなくても、学習データからモデルを構築する。機械学習は、データクレンジング、特徴の選択、トレーニング、テスト、検証などの問題解決プロセスにおいて、経験によって学習し、改善するさまざまなアルゴリズムを実装する。企業や政府機関は、不正行為の検知、推奨システム、顧客からの問合せ、医療、輸送、物流などのタスクに機械学習アルゴリズムを導入している。略語: ML

Reliability; 信頼性

学習されていない新しいデータが使用されたとしても、期待どおりに動作し、意図した機能を一貫して正確に実行することを保証する、AI システムの一属性。

Synthetic data; 合成データ

実際のデータの構造や統計的特性を模倣して類似させることができる、システムまたはモデルによって生成されたデータ。特に、現実世界のデータが限られていたり、入手できなかったり、または高度に機微な内容で利用できないような場合に、機械学習モデルのテストまたはトレーニングに使用されることが多い。

Training data; 学習（トレーニング）データ

学習データ内のパターンを見つけ、構造を識別し、結果を正確に予測できるようになるまで機械学習モデルをトレーニングするために使用される、データのサブセット。

【日本語版注記】

「学習データ」が、モデルが学習するために使用する全てのデータセットを指すのに対し、モデルが予測を学ぶための「正解」を提供するデータセットのことを「教師データ（Labeled data）」と呼ぶことがある。

Transparency; 透明性

AI が使用されているかどうか、そのモデルがどのように機能しているかといった説明などの情報開示を含む、AI システムに関する情報を利害関係者が利用できる程度。AI のアルゴリズムが機能し、意思決定を行う方法におけるオープン性、理解しやすさ、説明責任を意味する。

Trustworthy AI; 信頼できる AI

ほとんどの場合、責任ある AI や倫理的 AI という用語と同じ意味で使用されるが、これらはすべて、原則ベースの

AI ガバナンスと開発を指す。これには、セキュリティ、安全性、透明性、説明のしやすさ、説明責任、プライバシー、非差別／非バイアス（「バイアス」を参照）などの原則が含まれる。

Validation data; (妥当性) 検証データ

学習段階中に機械学習モデルの性能を評価するために使用されるデータのサブセット。検証データは、モデルのパラメーターを微調整し、テストデータを使用した最終評価の前に、オーバーフィットを防ぐために使用される。

注) 本用語集の公表または再掲載にあたっては、いかなる部分であっても IAPP から直接許可を得るべきである。

プロジェクトチーム

Barry Jennings, Director, Cloud + Commercial Business Planning Insights at Microsoft

Kathy Joe, Consultant to ESOMAR Professional Standards Committee

Judith Passingham, Chair of ESOMAR Professional Standards Committee

Ravinder Roopra, Head of Privacy at Kantar

Phil Sutcliffe, Managing Partner, Nexxt Intelligence

Facilitated by:

Paula Fernandez, Standards Programmes Coordinator at ESOMAR

2024 年 3 月 12 日 初版発行

(日本語版は 2024 年 4 月)

日本語版翻訳：一般社団法人 日本マーケティング・リサーチ協会 (JMRA)