（仮訳：12/23 版）

ESOMAR

GLOBAL RESEARCH BUSINESS NETWORK
APRC • EFAMRO • ARIA • AMRA

ESOMAR/GRBN Guideline When Processing Secondary Data for Research (Draft)

調査のための二次データ処理に関するガイドライン（ドラフト）

JMRA

2021 年 12 月

目　次

注）他のガイドラインには登場しない本文書独自の表現や記述と、相違部分を赤字で示している。

1　はじめに

市場・世論・社会調査及びデータ分析（以下「調査」）は、その歴史を通して、人々の行動、ニーズ、態度に関する情報やインサイトを提供し、商品やサービスの供給者、政府、個人、そして社会全体の意思決定に貢献してきた。その際、私たちは主に、調査参加者個々人との直接的なやり取りや観察を通じて収集されたデータに依拠してきた。

この 20 年ほどの間に、私たちはデジタル革命を経験してきた ― 情報の収集・保存・処理・分析の能力、グローバルなインターネット、ソーシャルメディア、モバイルテクノロジーの飛躍的な向上が見られており、人々の生活や仕事の方法が大きく変化している。その結果、すでにデジタル形式で利用可能となったデータへの依存度を高めることによって、調査は変貌を遂げつつある。リサーチャーの役割は、インタビュアーまたはデータ収集者からデータキュレーターへと進化し、データの整理と統合により重点を置くようになっている。調査とインサイトの機能は、一次データの収集と分析だけでなく、さまざまなソースからのデータの管理、合成、分析にまで拡張されており、多くの場合、新しい分析コンセプトと技術の使用法が進化している。その結果、リサーチャーが大規模なデータベースを組み上げて分析し、パターンを発見して強力な新しいインサイトを提供するという、調査に対する全く新しいアプローチが生まれた。

同時に、個人（以下「データ主体」）が自らの個人データがいつ収集され、どのような目的で、どのように使用されるかを決定できることの重要性に対する社会的関心が高まっており、責任を持ってそのデータを取り扱う方法について、明確で倫理的・専門的なガイダンスが緊急に必要とされている。

このような変化にもかかわらず、リサーチャーは意思決定者や他のデータ利用者に対して、データ処理や分析の詳細についてオープンで完全な透明性を保つという倫理的責任を持ち続けている。そのような透明性は、調査のユーザーがその品質を評価し、目的に適しているかどうかを判断する唯一の方法である。

2　目的と対象範囲

本ガイドラインでは、リサーチャーの倫理的責任について、彼らが所属する組織の種類にかかわらず、二次データ（既に存在するデータ）を利用する業務に従事する場合について述べる。このデータには、データ主体が企業や政府機関とやり取りする際に作成されるトランザクションデータベース、ソーシャルメディア・ネットワーク、シンジケートデータ、IoT から生成されるセンサーやスキャナーのデータ、その他多くの類似した種類のデータなど、幅広く多様な情報源から取得・収集されるデータが含まれる[1]。

本ガイドラインは主にリサーチャーを対象としているが、読者にはクライアントやその他のデータユーザーも含まれており、彼らが自らの責任を十分に認識し、確立された倫理的及び法的要求事項の下で、何が可能で何が不可能かについての期待値を設定することを意図している。

ここに記載されている要求事項及びベストプラクティスは、特定の国または地域の法的要求事項を反映するものではない。むしろ、ICC/ESOMAR の市場・世論・社会調査及びデータ分析に関する国際規範、既存の ESOMAR/GRBN のガイダンス文書、世界各国の市場調査協会の綱領とガイドライン類を補完するように設計されている。したがって、これらは単独で使用すべきではない。

[1] 受動的データ収集でも、リサーチャーがデータ主体と直接対話して行動を観察し記録するために同意を取得するような形態は含まれない。ESOMAR/GRBN「一次データ収集に携わるリサーチャーとクライアントのためのガイドライン」を参照。

この ESOMAR/GRBN の指針もまた、国内法に優先するものではない。国際プロジェクトに責任を負うリサーチャーは、本ガイドラインの規定を最低限の要求事項とし、法律または国が合意した基準に定められたその他の責任をすべて果たすべきである。これは法的助言ではなく、そのように解釈してはならない。リサーチャーの責任は、調査に影響を与える可能性のある、いかなる法律にも遅れずについていくことと、関係者全員がその要求事項を認識し、それに従うことに同意することである。

この文書では、必須の要求事項を特定するために「しなければならない」という語が使用されている。リサーチャーが守らなければならない原則や実践を表すときに、「しなければならない」という言葉を用いる。「望ましい」という言葉は、推奨事項を説明するときに使用される。この用法は、リサーチャーが調査の設計に応じて異なる方法で原則または調査活動を実施する際に、選択ができることを伝えるように意図されている。

3　用語の定義

本書の目的上、これらの用語は以下の特定の意味を持つ。

API（アプリケーションプログラミングインタフェース）

コンピュータ・プログラムが他のプログラムまたはコンポーネントと通信する際の基準となり、また、内部または外部でのアクセス／データ交換をサポートできる一連の定義。

Automated decision-making system（自動化意思決定システム）

人間が介入することなく反復的な運営面の決定を行う、定められたルールに基づいて稼働するシステム。

Children（子供）

調査に参加しようとする個人は、親、法定後見人、または責任ある大人から許可を得なければならない。子供の年齢の定義は国によって大きく異なり、各国内法及び自主規制規範によって定められている。国による定義がない場合、子供は 12 歳以下、「若者」は 13 歳から 17 歳と定義されている。
（日本語版注記：日本では JMRA 綱領で「中学生以下」と定義している）。

Client（顧客）

調査プロジェクトの全部または一部を依頼、委託、または申し込む、個人または組織。

Consent（同意）

個人データの収集及び処理に対して、本人の自由意思に基づいて与えられ、かつ、明確に示された合意。

Data analytics（データ分析）

調査目的のために、隠れたパターン、未知の相関関係、傾向、好み、及び他の有用な情報を明らかにするために、データセットを調べるプロセスを意味する。

Data provenance（データの出所）

1 つ 1 つのデータの起源と、それがデータベース間を移動する場合に追跡すること。

Data subject（データ主体）

その個人データが調査に使用されるすべての個人。

Deductive disclosure（演繹的特定）

クロス分析、少量サンプル、または他のデータ（クライアントの記録や公的機関の二次データなど）との組み合わせによって、あるデータ主体を推論的に特定すること。

Harm（危害）
有形的及び物質的な損害（身体的な傷害や金銭的損失など）、無形的または道徳的な損害（評判または信用の失墜など）、または私的な生活への過度の侵害を意味し、求められていない個人を標的としたマーケティング・メッセージを含む。

Non-research activity（非調査活動）
データを収集または分析された個人に対して、その個人の態度、意見または行動を変える目的で直接的な働きかけを行うこと。

Passive data（受動的データ）
個人の行動または態度を観察、測定または記録することにより、個人データを収集すること。

Personal data（個人データ：時には個人を特定できる情報、または PII と呼ばれる）
例えば、直接的な識別子（名前、特定の地理的位置、電話番号、画像、音声、ビデオ録画など）を参照することによって、または間接的に個人の身体的、生理学的、精神的、経済的、文化的若しくは社会的特徴を参照することによって、個人を特定するために使用することができる自然人に関するあらゆる情報。

Primary data（一次データ）
調査の目的のためにデータ主体から、またはデータ主体についてリサーチャーが収集したデータ。

Privacy（プライバシー）
個人が他からの侵害や干渉から自由であり、自分自身に関する情報を制御、編集、管理及び削除する能力を有し、そのような情報をどのように、どの程度他人に伝達するかを決定する能力を有することを前提とする個人の権利。

Privacy Impact Assessment（プライバシー影響評価：時に PIA または DPIA と呼ばれる）
データ主体のプライバシー・リスクを特定し、軽減するプロセス。

Profiling（プロファイリング）
調査以外の目的でデータ主体に直接的な働きかけを行うために、データ主体の業務上のパフォーマンス、経済状況、健康状態、個人の嗜好、関心、信頼性、所在地、行動を分析または予測する目的で、個人データを収集及び処理すること。

Research（リサーチ：すべての形態の市場・世論・社会調査及びデータ分析を含む）
個人や組織に関する情報の体系的な収集と解釈。社会科学、行動科学、データ科学を応用した統計的・分析的方法と技術を用いて、インサイトを生み出し、企業、政府、非営利団体、一般市民による意思決定を支援する。

Researcher（リサーチャー）
調査に関するコンサルタントとして活動する個人または組織。クライアントの組織及び使用される二次契約業者で働く者を含む。

Secondary data（二次データ）
すでに収集され、別の情報ソースから入手可能なデータ。

Segmentation（セグメンテーション）
広範な対象母集団を、共通のニーズ、関心、優先事項を持っているか、持っていると思われる個人または組織のサブセットまたはグループに分割し、それらを相互作用するための戦略を設計し実行することを目的とした

分析技術。セグメンテーションはプロファイリングとは異なり、個々のデータ主体ではなく、共通の特徴を持つ明確に定義された人々のグループに焦点を当てている。

Sensitive data（機微なデータ）（国によっては「特殊カテゴリーデータ」とも呼ばれる）
個人や組織のプライバシーまたはセキュリティを保護するために、現地の法律で許可されていないアクセスから可能な限り高いレベルで保護することが要求されている特定の種類の個人データであり、処理前にデータ主体からの追加の明示的な許可が必要になる場合がある。機微なデータの指定は管轄区域によって異なり、データ主体の人種または民族的出自、健康記録、生体情報及び遺伝子データ、性的指向または性的習慣、犯罪記録、政治的見解、労働組合への加入、宗教的または哲学的信念などを含むが、これらに限定されない。また、所在地、財務情報、規制薬物またはアルコールの使用などの違法行為など、その他の種類のデータ（必ずしも法的に定義されているわけではない）を含む。

Vulnerable individuals（保護を要する人々）
認知障害またはコミュニケーション障害を有する人を含め、得られた情報に基づく意思決定を自発的に行う能力が限られている人々。

Web scraping（ウェブスクレーピング）（時にクローリングまたはスパイダーとも呼ばれる）
Web サイトからデータを抽出するためのソフトウェアを使用すること。

4　主要な原則

市場・世論・社会調査及びデータ分析の長い歴史を通して、リサーチャーは、個人データがいつ、どのように収集され、使用されるかを決定する固有の権利が個々のデータ主体にあることを認識してきた。この目的のために、私たちの業務は 3 つの最も重要な原則によって統治されてきた。

- 調査の目的のためにデータ主体から個人データを収集する場合、リサーチャーは、収集する予定の情報、それが収集される目的、それが誰と共有される可能性があるか、どのような形式であるかについて透明性を確保しなければならない。
- リサーチャーは、調査で収集され、使用される個人データが、不正なアクセスまたは使用から完全に保護され、データ主体の同意なしには開示されないことを確実にしなければならない。
- リサーチャーは、常に倫理的に行動し、適用されるすべての法律及び規制を遵守し、データ主体に危害を及ぼしたり、市場・世論・社会調査及びデータ分析の評判を傷つけたりするようなことをしてはならない。

これらの原則[2] は、リサーチャーがデータを依拠している一般市民と、より良いビジネス上の意思決定を支援するために調査を発注したクライアントとの信頼の基盤を形成する。これらの原則は、私たちの長い歴史の中でいつでもそうであったように、今日においても重要である。

二次データは、リサーチャーが収集条件をあまりコントロールできず、使用される同意メカニズムが十分に強固ではないか、全く存在しない可能性のある変化する環境に適応することをリサーチャーに要求する。同時に、二次データに含まれ、調査で使用される個人データが法的根拠なしに開示されないこと、個人データの使用が危害を及ぼしたり、その他の悪影響をもたらしたりしないことを確実にしなければならない。

---

[2] 経済協力開発機構 (OECD) も同様のプライバシー原則を支持しており、世界中の多くの既存及び新興のプライバシー及びデータ保護法に反映されたフレームワークを構成している。詳細は、OECD Privacy Framework を参照。

データ主体に対する責任

5　調査の設計

リサーチャーは、個人データを依拠するデータ主体に対して倫理的責任を持ち、自主規制部門のメンバーとして設計段階からこれらの義務を果たす。一部のガイダンスは、調査が実施される国の規制要件及びデータ保護要件によって提供されることがある。しかしながら、規制要件は国によってかなりのバラツキがあり、他国よりも制限が厳しい国もあれば、データ保護法が全くない国もある。リサーチャーは、データを収集または処理する国の法律を認識し、遵守しなければならないが、倫理的責任を果たすためには、単純に現地の法律を遵守する以上のことを求められる。そのための効果的な方法の 1 つは、「プライバシー・バイ・デザイン」と呼ばれる手法を使用することである。

5.1　プライバシー・バイ・デザイン

プライバシー・バイ・デザインの本質は、あらかじめ、予防的に、最初から最後までのプライバシー保護がデフォルトで設定されていることを重視するプロセスの実装である。ここで適用されているように、それは 3 つの主要な構成要素を持っている。（a）明確に定義されたプライバシー原則の基礎、（b）特定のプロジェクト設計におけるプライバシー・リスクを評価するためのプロセス（例えば、プライバシー影響評価）、（c）これらのリスクを軽減するための、情報セキュリティの実践とプライバシー保護のアプローチ、方針、及び手続きのインフラストラクチャである。

プライバシー・バイ・デザインと、グローバルなプライバシーフレームワークの基本原則の 1 つは、データの最小化であり、個人データの収集を、特定の目的を達成するために直接関連し、必要なものに限定する慣行と大まかに定義される。時間と費用を現実的に考慮すると、一次データ収集におけるデータの最小化が促進される。二次データを扱う場合、しばしば膨大な量の利用可能なデータとそれを処理するための計算能力が、リサーチャーを「すべてのデータ」の収集に集中させ、どのデータが分析段階に関連するかの判断を置き去りにすることがある。その結果として、大量の二次データを処理する場合に、十分に考え抜かれた堅牢なデータ保護対策の必要性が、大幅に高まることになる。厳格なプライバシー影響評価は、そのための不可欠なツールである。

5.2　プライバシー影響評価

慎重に実施されるプライバシー影響評価 - PIA（Data Protection Impact Assessment：DPIA とも呼ばれる）は、ある特定の調査設計が、個人データが調査に使用された結果として悪影響や有害な結果を経験しないように、データ主体の個人データとプライバシーの必要な保護を含むことを保証する。簡単に言えば、PIA とは、プロジェクトのライフサイクルを通して、データ主体の個人データとプライバシーに対するリスクを体系的に特定し、軽減するプロセスである。通常、次の 4 つの手順がある。

1. プロジェクト及び関連する全ての組織を通じた、計画された情報の流れを図示する。
2. リスクを特定し、その重大性及び発生可能性を評価する。
3. 特定されたリスクを軽減するソリューションを開発及び評価する。
4. リスク軽減ソリューションを組織のプロセスと計画に統合する。

5.3　追加ガイダンス

PIA のより詳細な取り扱いについては、ESOMAR／GRBN の「注意義務に関するガイドライン：調査対象者保護のために」を参照のこと。さらに、ESOMAR の「Data Protection Checklist（データ保護チェックリスト）」は、組織の情報セキュリティ・インフラストラクチャーと実践上のギャップを特定し、ソリューションを開発するための段階的な評価プロセスを提供している。リサーチャーは、PIA のリスク軽減フェーズの一環として、それを参照することが望まれる。

6　個人データ処理の法的根拠の確立

世界中のデータ保護フレームワークでは、個人データを収集または処理する前に、あらゆる種類の個人及び組織が明確で説得力のある法的根拠を確立する必要性がますます高まっている。これらの要求事項は、リサーチャーにも適用される。法的要求事項が存在しない国や地域においても、リサーチャーはデータ主体のプライバシーと権利を尊重する責任がある。そのため、個人データを収集または処理するための何らかの法的根拠を確立する必要がある。

この要求事項には、処理対象のデータの所有者を特定し、データを処理するための許可を得る必要があることが暗黙のうちに含まれている。リサーチャーは、データ保持者の理解と同意なしに、ウェブサイトや他のオンラインソースから個人データにアクセスしたり、取得したりしてはならない。

6.1　データの出所の確認

個人データを含む二次データソースにアクセスする前に、リサーチャーはまず、個々のデータ項目の出所、すなわち、データの出所とその後の処理状況について可能な限り詳細に確認しなければならない。複数の情報源から構築されたデータベースを使用する場合、すでに多くのマージ、リンク、変換または集約の手順が実行されている可能性があるため、これは困難な場合がある。難易度は、そのデータが一次データなのか、二次、三次のいずれであるかによって異なる。例えば、一次及び二次のデータを扱う場合には、データの保有者を識別し、収集の状況を確認することは容易であることが多い。しかし、サードパーティのデータ (一般的にはマルチソース) を扱う場合、出所を確定することさえ大変な作業になることがある。例えば、データブローカーは通常、数十の情報源から個々の消費者のプロファイルを作成するため、データ収集時にどのようなことがデータ主体に伝えられたか、また、その使用にどのような制限が課されたかを検証することは困難である。

そのための 1 つの簡単な方法は、データソースごとに、収集時にデータ主体に提供される利用規約 (ToU) 、プライバシー通知、またはその他の類似文書を入手し、レビューすることである。リサーチャーは、どのようにして、どのような条件で、どのような目的でデータが収集されたかを特定する情報によって十分に裏付けられた個人データを含む、または構成する二次データソースのみを使用しなければならない。とりわけ、リサーチャーは、個人データが合法的かつデータ主体の理解を得て収集されたものであることを検証しなければならない。その場合にのみ、リサーチャーはデータを調査目的で処理できるかどうかを判断することができる。

6.2　特定の根拠の選択

個人データを処理する法的根拠を確立するための要求事項は、世界中でますます一般的になってきているが、利用可能な根拠、適格性の認定方法、及びどのような特定のデータ収集また処理活動が許可されているかとい

う点で、行政区域によって大きな違いがあることが多い。したがって、リサーチャーは、収集される個人データに適用されるすべての行政区域内の要求事項を十分に理解し、関連する法律を遵守することを確実にしなければならない。

6.2.1 通知と同意

一次データ収集に従事する際、リサーチャーは一般に、あらゆる形態の個人データを収集・処理する前に、データ主体からの同意に頼ってきた。これには、収集を計画している情報、収集の目的、保護の方法、共有する相手、及びその形式に関する透明性の確保が含まれる。

二次データを扱う場合、利用規約（ToU）で表現されている同意手法の厳密さは大きく異なる。古典的な調査の同意プロセスと同じ要素を多く含むものもあれば、重要な欠落要素があるものもある。さらに、データ主体は、同意を示す前に注意深く読まずに ToU に同意した可能性がある。

リサーチャーが処理の根拠として同意に依拠しようとする場合は、次のことを判断するための十分な情報がなければならない。

- データは、偽りなく、またはデータ主体にとって明白で、合理的に識別でき、予測可能な方法で、合法的かつ透明性をもって収集された。
- データ主体は、個人データを共有することに同意を求められた。
- データが使用される目的が明確に特定されていた。
- 収集時に提供された ToU またはプライバシー通知のいずれにおいても、調査のためのデータ使用は明確に除外されていない。
- 個々のデータ主体から、収集時に説明された目的以外にそのデータを使用しない旨の要請があった場合は、その要請を尊重する。

これら 5 つの条件のいずれかを満たさない場合、リサーチャーは他の根拠を検討する必要がある。

6.2.2 正当な利益

正当な利益は、個人データを処理するために使用することができる同意に代わる根拠を提供する。正当な利益は、データ主体が合理的に期待する方法で個人データが使用されており、その処理がプライバシーに重大な影響を与える可能性が低い場合に、処理に適した根拠となり得る。

正当な利益とは、すべての利害関係者 (データ主体、データ保有者、クライアントまたはその他のエンドユーザー、さらには社会全体) の利益のバランスを明示的に考慮することである。個々の利害関係者は、新しい、潜在的に有用なインサイトを発見するためにデータを処理することに様々な利害関係を有し、それらの利害関係は対立する可能性がある。その結果、リサーチャーは、これらの競合する利益と、データ主体の利益に最大の重点を置くこととのバランスを取らなければならない。これは、特に厳格な PIA と、データ主体に対する潜在的な危害の可能性を防ぐための強力なプライバシー及びデータ保護措置の必要性を意味する。

正当な利益を使用できるかどうかを判断する際には、リサーチャーは、データ主体の基本的な権利及び自由よりも、自らの利益またはクライアントの利益が優先されないようにしなければならない。正当な利益を処理の根拠として使用することを検討する場合、リサーチャーはこれらの基準に対応する 3 段階のアプローチに従って、文書化しなければならない。

- 目的 ― 正当な利益が追求されているか？

- 必要性 ― 目的を達成するために、その処理が必要か？
- バランス ― データ主体の権利と利益が、利害関係者の利益に優先されているか？

正当な利益の利用を検討する際に、データ主体の利益を考慮し、比較検討するプロセスは、例えば、正当な利益評価として何らかの方法で文書化されなければならない。正当な利益が機微なデータの（「特殊カテゴリーの個人データ」と呼ばれることもある）処理や、自動化された意思決定のために使用されてはならない。

6.2.3 互換性のある目的

互換性のある目的はまた、別の法的根拠を提供する。二次データの使用は、収集時にデータ主体に提示されたものから目的の変更を伴うことが多い。行政区域によっては、目的の変更により、データを収集された人の再同意が必要となる場合がある。これは困難であり、リサーチャーが新しい目的の詳細をデータ主体に連絡する必要が生じる場合がある。それ以外には、データ管理者のウェブサイトに通知を掲載するだけで、データ主体に同意を取り消す機会を提供すればよい場合もある。

目的の変更に際して、新しい目的への同意を取得する必要がない場合もある。1つは、「法律に基づいて」という単純なものである。新しい目的が類似している、すなわち互換性がある目的は、もう1つの例である。

互換性のある目的を確立するためには、データ収集時の当初の目的と新たな目的との関係を慎重に検討し、データの将来的な利用の可能性に関するデータ主体の合理的な期待とのバランスをとる必要がある。また、公正な処理を確保し、データ主体のプライバシーへの影響を制限するために、適切な緩和措置が講じられていることを前提としている。例えば、オンライン小売業者は通常、顧客の購入行動、支払い方法、プロモーションへの対応、及び製品の配送やサポートに必要なその他の個人データに関する情報を収集する。小売業者は、どの製品をどの程度の価格で提供するか、どのように販促するのが最適かなどの理解を深めるために、そのデータを使用すると想定するのが妥当であろう。この場合、データの使用は収集の当初の目的に適合する。これには、対象を絞ったマーケティング・メッセージを、それを受け取ることを選択した個々の顧客に配信することも含まれる。

一部の行政区域では、統計調査は互換性のある目的と考えられている。このような場合であっても、リサーチャーは本ガイドラインに記載されているプライバシー保護措置を遵守しなければならない。

6.2.4 契約

契約もまた、個人データを処理するための根拠として使用することができる。リサーチャーは、データ主体に対する契約上の義務を履行するためにデータ主体の個人データを処理する必要がある場合に、この根拠を利用することができる。これは、調査の文脈では適用が限定されているが、アクセスパネルを管理し、運用する際には適用可能である。

6.2.5 公共の職務と利益

公共の利益、または公的な職務の遂行のために必要な個人データの処理は、リサーチャーが考慮することのできるもう1つの根拠である。この根拠を用いるための条件は厳密に定義され、国によって異なる傾向がある。結果として、公共部門による調査や、公共の利益となることを明確に示すことのできる民間部門の調査で主に使用されている。

7 データセキュリティ

リサーチャーは、データ処理中に（a）データ主体のプライバシーを完全に保護し、（b）処理及び分析中にエラーが発生しないようにすることを確実にしなければならない。どちらの場合も、リサーチャーはこれらの課題を達成するために設計された一連の手順と基準を整備しておかなければならない。

7.1　プライバシー保護

ESOMAR の「データ保護チェックリスト」は、個人データの不注意な漏洩や消失を防ぐために設計された技術、基準、及びプロセスのインフラストラクチャへのロードマップを提供する。リサーチャーは、このツールをプライバシー保護プログラムの評価ツールとして使用し、ギャップを特定して解決策を策定することが望ましい。

重要な関心事は、個人情報がクライアントに開示されないことである。リサーチャーは、適用されるプライバシーに関する法律または規制でより高い要件が規定されていない限り、以下の条件の下でのみ、データ主体の個人情報をクライアントに伝達することができる。

- データ主体が明示的な同意を与え、
- 目的が調査のためのみである。

さらに、リサーチャーは、上記の条件が満たされない限り、クライアントがデータ主体を再識別しようとしない、という書面による保証をクライアントから得ることが不可欠である。詳細については、ESOMAR／GRBN の「注意義務に関するガイドライン：調査対象者保護のために」を参照のこと。

リサーチャーはまた、二次契約業者と共有する個人データが二次契約業務の遂行に必要なものに限定されていること、及び二次契約業者がそのデータを保護するために必要な情報セキュリティ手順を備えていることを確実にしなければならない。データ保護に関する二次契約業者の責任は、明確に文書化され、同意されなければならない。

7.2　文書化

リサーチャーは、データクリーニング、他のデータソースとの統合、重み付け、補定（使用している場合）、実施した特定の分析など、それぞれの具体的な処理手順を完全に文書化しなければならない。文書は、データ利用者が、データ処理の過程でどのようにデータが変更された可能性があるかを理解できるように、十分に具体的であることが望ましい。詳細は以下の 8.3 項を参照のこと。

クライアント及び他のデータ使用者に対する責任

8　透明性

8.1　プロジェクト設計

リサーチャーは、提案され、契約として合意された目的、仕様、品質を満たすように調査を設計しなければならない。リサーチャーは、調査の実施方法について最初から最後まで透明でなければならない。この情報は通常、提案の段階でクライアントに伝えられ、作業の進行に応じて変更される。ISO 規格「ISO 20252:2019 市場・世論・社会調査及びインサイト・データ分析－用語及びサービス要求事項」は、提案段階でクライアントや他のデータ利用者に開示し、調査の進展に合わせて更新すべきプロジェクト設計機能の詳細なリストを提供する。実施する調査や分析に使用される特定のデータの完全な透明性を確保するために、記載された要求事項を遵守することが望ましい。

8.2　二次契約

リサーチャーは、業務を開始する前に、業務の一部がリサーチャーの組織外に委託される予定であることをクライアントに通知しなければならない。要請がある場合、クライアントにそのような二次契約業者の身元を知らせなければならない。

リサーチャーはまた、二次契約先と共有する個人データが二次契約業務の遂行に必要なものに限定されること、二次契約先がデータを保護するために必要なデータセキュリティ手順を整備していること、二次契約先のデータ保護責任が明確に文書化され、同意されていることを確実にすることが求められている。

8.3　文書化

調査が価値を持つためには、それが使用され、実行されなければならず、そしてデータ利用者がデータの妥当性、実施された分析、及びその結果の正確さを受け入れたときにのみ実行される。リサーチャーは、クリーニング、他のデータソースとの統合、重み付け、補定 (使用する場合) 、実施された特定の分析など、それぞれの具体的な処理および分析ステップを完全に文書化しなければならない。文書は、データ利用者が分析の過程でデータがどのように変更された可能性があるかを理解できるように、十分に具体的であることが望ましい。

アンケートやフォーカスグループなどの一次データ収集を使用する場合、リサーチャーやクライアントがほぼ同様に理解している、信頼できる使い慣れた広範な測定基準と、測定に対する規律あるアプローチがある。ISO 20252 に規定された要求事項の透明性のレベルは、クライアントや他のデータユーザーが、従来の一次調査の設計目的に対する妥当性や適合性について、十分な情報に基づいた判断ができるようになることを確実にする。これらの要求事項は、従来の分析を二次データで使用する場合にも適用される。リサーチャーが調査の最後にクライアントと共有しなければならないことの詳細については、一次データに関する ESOMAR／GRBN のガイドラインを参照のこと。

機械学習のような新しいアルゴリズム分析の使用が増えていることは、新たな課題を提起している。これらの技術はしばしば「不透明」または「ブラックボックス」と表現されるが、従来の測定基準の多くは依然として適用可能である。

最低限、リサーチャーは以下を文書化しなければならない。

- 調査に資金を提供した組織、調査を実施した組織、および使用した二次契約業者の名称；
- 調査の目的；
- 対象母集団の定義；
- 使用したデータソースとその理由；
- 含まれているデータ項目とその情報源のリスト、補定を実施した場合には使用した方法；
- 統計解析の方法 (あてはまる場合) ；
- 複数の情報源からのデータを組み合わせた場合、使用した技術とその精度の評価方法；
- 必要に応じて、データのエディティングまたはクリーニングに使用された方法；
- データの欠落レベルの評価；
- 分析の信頼性、精度、妥当性を評価する頻度とプロセス。

リサーチャーはまた、以下も考慮することが望ましい。

- 実施する調査の複雑さと性質に関連するリスクを特定し、評価し、対処する；
- 調査結果の妥当性に影響を及ぼす実質的な限界についての記述を含める；
- 利害関係者 (例：クライアント、コミュニティ、規制当局) のニーズと期待を特定し、彼らの要求事項が考慮されていることを保証する；
- 監査と再現を可能にするための明確性、透明性、識別性、トレーサビリティを提供する；
- バイアスが生じることが知られているか、疑われている方法について慎重に文書化する。

8.4　機械学習（マシンラーニング）

機械学習を使用する場合、追加的な一連のレポーティング要求事項がある。機械学習の演習の典型的な目標は、新しい入力データを分類できるモデルを構築することであり、しばしば予測を行うことである。これらの分類または予測の精度を評価するために一般に受け入れられている方法は、一連のよく設計されたテストサンプルを開発して提出することであり、そこから精度測定基準を計算して評価することができる。これらの指標には、分類精度、対数損失、混同行列、曲線下面積、F1 スコア、及び平均絶対誤差が含まれるが、これらに限定されない。

一般市民に対する責任

9　結果の公表

クライアントが調査プロジェクトの結果を公表しようとする場合、クライアントとリサーチャーの両者が、公表された結果が誤解を招くものでないことを確実にする責任を負う。そのために、クライアントは調査結果の公表の形式と内容についてリサーチャーと協議することが強く推奨される。リサーチャーはまた、要請された場合に、公表された知見の妥当性を評価するのに十分な技術情報を提供する準備ができていなければならない。

リサーチャーは、市場調査プロジェクトから得られた結論がデータによって適切に裏付けられていない限り、その結論の公表に自らの名前を付け加えてはならない。

10 参考文献

11 プロジェクトチーム

＜日本語版作成＞

一般社団法人　日本マーケティング・リサーチ協会