データプライバシー・倫理・保護
—2030 年アジェンダ目標のビッグデータに関するガイダンスノート
—1

本ガイダンスノートの目的

- ビッグデータの使い方に関わっている国連開発グループに向けたデータプライバシー・倫理・保護の一般的なガイダンスを示し、2030 年アジェンダ目標を支援するためのプログラムを運用実装するために、UNDG のメンバーとの共有を行なった。
- 本ガイダンスノートでは次の 3 つを行なっている。
  ① SDGs の達成を目標としたビッグデータの運用を支援するために、UNDG 内で共通原則の設定
  ② 人権を考慮に入れたリスクマネジメントツールとしての活用
  ③ 民間部門からのデータの獲得、保持、使用、品質管理に関する原則の設定
- 「データ革命」は SDGs の達成を促進するものであると考えられていた。それは、進捗状況を見るだけでなく、あらゆるレベルのステークホルダーがエビデンスに基づいたポリシーを進め社会的弱者との関わりを持てるように働きかけるからである。
- SDGs の進捗状況確認の支援や全ての人に関するデータを収集できているかということを確認するために、品質が高くアクセスが可能、タイムリーで信頼性の高いデータが求められており、こうしたデータが意思決定において重要であると 2030 アジェンダは主張している。
- データ利用の話が進む一方で、ビッグデータの取扱や処理におけるリスクに関する法的懸念も存在している。2030 アジェンダの達成のために、ビッグデータ利用の安全性や責任が明示されたフレームワークを作成することが求められている。
- この文書で述べられているガイダンスは、国連総会決議 45/95 で採択された、コンピュータ化された個人データファイルの規制に関する国連ガイドラインを元に作成され、現存する国際的な方法（international instruments）UNDG メンバー

[1] United Nations Development Group, Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda, https://unsdg.un.org/sites/default/files/UNDG_BigData_final_web.pdf, last visited June 18, 2021.

本資料は，JST，RISTEX，JJPMJRX19B5 の支援を受けた研究成果の一部です。

のポリシー等も考慮に入れている。ガイダンスノートは時間をかけて制作され、その中心的価値を反映している。

- この文書ではプライバシーの権利が基本的人権であることを確認してデータの社会的価値を認識し、UNDG とそのパートナーとともに説明責任や透明性が担保された責任あるデータ利用の実践に関する一般的な枠組みを設置することを目標としている。
- ガイダンスノートは法的文書ではなく、最低限の自己制御を提唱することしかしない。そのため、実装組織によって拡張・詳細化がなされる場合がある。
- ビッグデータ利用により生じる利益やリスクを考慮し、個人や団体のプライバシーについて考える。また、標準的モラルや倫理的実践を考慮に入れ、ビッグデータの利用の際の前後関係の重要性を理解する。
- このガイダンスノートは、より詳細な運用ガイドラインを通じて実行されることが望ましい。このガイドラインは UNDG メンバー組織の権利や規則、ルールなどの実行についての説明を行なっている。なお、このガイダンスノートの実施及び遵守については、必要に応じて指定された法律、倫理、プライバシー、セキュリティの専門家に相談することが望ましい。ガイダンスノートの実施及び遵守を監視するメカニズムを構築するために組織を実装することが推奨される。
- 今後のテクノロジー分野の変化に応じてこのガイダンスノートは更新を行なっていく。

9 つの原理

① 合法的で正当公正な利用

- データアクセスや分析などにおいて、国連の規則やデータ保護法等を遵守する必要がある。
- リスクや害、利益の査定を行うことにより、個人や団体に不利益が被るような方法でデータ利用を避けなければならない。
- ビッグデータには個人情報やセンシティブなデータが含まれているので、その利用のためには、UNDG メンバー組織の規則・ルール・ポリシーに関する次の 4 要件のうち 1 つ、ないし複数を満たすことが求められる。

  I. データが利用される人物の同意

  II. 法律の遵守

  III. 国際的組織義務の促進

  IV. 個人や団体の利益を守るための正当なニーズ

② 使用目的や使用制限、目的の互換性

- データ使用目的に応じてデータは使用されなければならない。その使用目的はできるだけ狭くかつ正確に定義されていなければならない。データアクセスの申請は特定の目的に合わせて調整される必要がある。
- 当初の目的と異なるデータ利用を行う際には、その互換性（変更によりどのような影響が個人・団体に生じるかなど）を明示しなければならない。
- データアクセスの目的は、アクセスする際に明示されなければならない。

③ リスク低減とリスクや危害、利益の評価

- リスクや危害、利益の評価やリスク低減は、新しいデータ利用やその利用方法に変更が加えられる前に実施されなければならない。
- この評価を行う際には、データ利用により発生する副次的な危害が発生する可能性があるため、データ利用のコンテクストを考慮に入れる必要がある。また、データ利用がもたらす個人・団体への影響も考慮に入れる必要がある。
- 危害評価において次の 3 つが重要となる。

  I. 危害発生の可能性

  II. 危害の大きさ

  III. 危害の深刻さ

- 評価の際にはデータ利用者やデータの持ち主のデジタルリテラシーを考慮に入れる必要があり、可能であれば、様々な専門家から構成されるグループが評価を行うことが望ましい。
- センシティブなデータほど危害のリスクは大きくなるので、より厳しい保護

方法を採る必用がある。また、こうしたデータの利用に関する意思決定は、可能であれば、関係するグループや代表者との協議を行う必要がある。

- セクション 3 で記載されているようなデータ侵害やセキュリティシステムに関するリスクを考慮に入れることは重要である。
- データ利用は比例の原則（the principle of proportionality）に則り、リスクや危害が利益を上回るようなことがあってはならない。また、相互に関連する個人の権利に対するデータの影響を評価することも推奨される。
- これらのような評価査定は、データ利用が正当なものであるかを判断する材料となる。

④ センシティブなデータとコンテクスト

- データを保持し利用している間、データ保護に関して厳しい基準を設けることが求められる。コンテクストによってはあるデータがセンシティブなデータとなり、データ分析がもたらす個人・団体への影響に変化を与える可能性がある。

⑤ データセキュリティ

- データが外部に漏れることなどを防ぐために、テクノロジーやコストなどを考慮に入れたデータセキュリティは重要な役割を果たしている。
- プライバシー・バイ・デザイン（Privacy by Design）の基本原則とプライバシー強化テクノロジーを導入することは強く推奨される。
- プライバシー侵害の危険性を抑えるために個人情報の匿名化は施されるべきであり、必要ならば、UNDG メンバー組織は第三者データプロバイダーによって匿名化されたデータを使用することを選択肢に入れることが勧められる。
- 正当な理由のない匿名化データの再識別かを防ぐために、データを匿名化処理した人物によってそのデータが分析・利用されることを禁じることが望ましい。
- プライバシー保護のためにとった方法により、特定のデータ使用目的において不釣り合いが生じることがないように確認することが重要である。なおその方法は、データから生じる利益を最大化し、その利用目的を達成することができるようなものとするべきである。
- データへのアクセスはデータプライバシーやセキュリティに関するトレーニングを受けたものに限定される冪である。また、データ利用の前には、セキュリティシステムの脆弱性を評価する必要がある。
- データセキュリティの方法は、データ利用のリスクや危害、利益に照らして評価する必要がある。

- データセキュリティシステムの脆弱性から発生しうるリスクを考える際、データの流出や違反をもたらす原因を考えることは重要である。このデータの流出や違反を引き起こす人物（団体）として、次の 3 つが挙げられる。
  I. データへのアクセスが認められている人物
  II. データアクセス権を保持、または要求している「善の」第三者組織や、データの不正利用を目的としてアクセスしようとしている「悪の」第三者組織
  III. 素性の知れない第三者組織
- クラウドサービスを利用する際には、それに伴うリスクや危害を考慮する必要がある。

⑥ データ保持と最小化

- 保持するデータは、その利用方法を参考に、目的を達成するために必要な最小限の量に留める必要がある。
- データ保持の際にはリスクや危害、利益を考慮して行わねばならず、データの削除の際には適切な処置が施されなければならない。

⑦ データの質

- データが関係する活動は設計・実行・報告・文書化をすることが必要である。より具体的には、合理的に可能な範囲で、データの正確性、関連性、十分性、完全性、完全性、使いやすさ、有効性、一貫性を検証し、最新の状態に保つ必要がある。
- 意思決定に質の低いデータを利用することでリスクが発生する可能性があるので、データの質は注意深く考える必要がある。
- バイアスによる悪影響を防ぐためにデータの質の評価をしなければならない。
- 個人や組織に影響を及ぼす可能性のある意思決定をするためにデータを分析する際、自動データ処理機能を使うことは控えるべきである。
- データの質の評価はその利用中に定期的に行われるべきである。また、古いデータの消去やアップデートをコンスタントに行う内部システムを構築することが重要である。

⑧ オープンデータや透明性、説明責任

- 関連法案との一貫性を確認するために、適切なガバナンスや説明責任メカニズムを構築する必要がある。
- 説明責任を実施する際の重要な要素として、透明性がある。データの使用法について透明性を担保は、そうすることによる危害が利益を上回る場合を除き推奨される。

- オープンデータはイノベーション、透明性、説明責任において重要な役割を果たす。しかし、個人情報の公開は避けられるべきであり、仮に公開する場合には、セクション 3 で述べられたようなリスクや危害の評価を行う必要がある。
- リスクや危害の評価は説明責任メカニズムにおける重要な要素であり、コンプライアンスを監視するためにどのガバナンスメカニズムが必要であるかを決定する際に役に立つ。また、この評価はどの程度データを公開するかや透明性を決定する際にも活用することができる。

⑨ 第三者コラボレーターに対する適切な注意

- 第三者コラボレータがデータを使用する場合、関連法案や国連の第三者コラボレーターがデータを使用する場合、関連法案、国連の義務、規制、規則、およびポリシーとガイダンスノートに従う必要がある。
- 第三者コラボレーターのデータ利用方法を評価するために適切なプロセスを踏むことが推奨される。
- 第三者コラボレーターから提供されたデータに安全にアクセスできることが確認できるように、データへのアクセスやデータ処理のパラメータを概説する法的拘束力に対して合意がなされることが望ましい。

定義とメモ

- データの集計（Aggregation of Data）
  - この文書におけるデータの集計とは、データから個人が特定されないようなフォーマットに個人データを結びつけるプロセスのことを意味する。集計されたデータは分析や統計を目的として利用される。
  - また、the UN Archives and Records Management Section は「グループやシリーズの編成がなされ集計されたデータ」と定義し、IOM Data Protection Manual では、「個人データから編集される可能性はあるものの、個々のケースを特定できないようにグループ化された情報（要約統計量など）」と定義されている。
- 適切な同意（Adequate Consent）
  - ある同意は情報化・文書化などを通じて初めて「適切な同意」となり、これはデータの収集前や元々の目的と異なるデータ利用が必要となった際になされる必要がある。
  - 形成された同意を情報化するために、データの使い方を通知に含めることが望ましい。即時的な同意は適切な同意に該当しない場合があるので、データによるリスクや危害、利益の比例性の査定を考慮することが重要である。
  - データ提供者はその同意の撤回やデータの使用に反対することができる。また、データ提供者が第三者である場合、その団体が適切な同意をしたか、或いはデータ共有に正当な根拠を持っているかを確認することは推奨される。
  - データの再利用の合意を図ることはしばしば難しいため、データ専門家はそのデータ利用がもたらす最善の利益について考える必要がある。そして合意が無い状態で利用を進めることを決定した場合、リスクや危害、利益を再評価することが求められる。
- ビッグデータ（Big Data）
  - UN Global Pulse はビッグデータを「構成・非構成化された大規模なデータで、従来のデータベースやソフトウェア技術では処理が難しいもの」と定義している。また、ビッグデータの特徴として、3V（more volume, more variety, higher rates of velocity）を掲げている。
  - ビッグデータは様々な種類があるが、この文書はビッグデータは様々な種類があるが、この文書はリアルタイムで民間部門に収集され、意思決定プロセスに影響を与える人間の行動を観測するために利用されるデータに着目している。
  - The International Telecommunication Union (ITU)はビッグデータを「リ

アルタイムの制約の下で、異なる特性を持つ広範なデータセットの収集、保存、管理、分析、および視覚化を可能にするためのパラダイム」と定義されている。

- データ匿名化 (De-Identification (Anonymization) of Data)
  - この文書においてデータの匿名化とは、「個人情報を匿名化データに変換するために合理的な方法を使用する過程」を指す。匿名化には様々な方法が存在する。
  - データの匿名化により全ての識別子が除去される場合があるが、それでも個人や団体を結びつける場合があるので、匿名化されていない個人情報と同レベルの保護を施す必要がある。
- デジタルリテラシー (Digital Literacy)
  - この文書においてデジタルリテラシーとは、「人々が扱い共有しているデータの認識方法」を指す。デジタルリテラシーはデータ利用者と提供者のどちらにも関わっている。
- 暗号化 (Encryption)
  - この文書において暗号化とは、「電子データを暗号コードに変更するセキュリティ手続きで、元のコードまたは暗号システムの助けなしに理解できないようにするために、コードまたは暗号システムのいずれかの方法を用いている手続き」を指す。
- 個人からなる団体 (Group(s) of Individual)
  - この文書では、the Office of the United Nations High Commissioner for Human Rights(OHCHR)が示す目には見えない「法的」グループも含めている。
- マスキング (Masking)
  - マスキングとは、ソーシャルメディアから収集した元の個人情報を匿名化する技術で、個人または団体を追跡したり関連付けたりできない程度に変更される。
- 個人情報 (Personal Information)
  - この文書において個人情報とは、「個人を特定できるデータ」を指すが、たくさんの地域や団体によって定義付けがされている。
  - 個人情報の公開・非公開はデータの所有者に一任されている。近年、ネットに個人情報の公開は普通になってきているが、これにより、公開したデータに関わる人々が危害を被る場合がある。
- プライバシー (Privacy)

- The Special Rapporteur to the Human Rights Council の報告書においてプライバシーとは、「個人が自律的な発達、相互作用および自由の領域、他者との相互作用の有無にかかわらず、国家の介入および他の個人による過度の介入のない『私的領域』を持つべきであるという推定」を指すが、国際的に一致している定義はない。

- プライバシー・バイ・デザイン（Privacy by Design）
  - Privacy by Design とは、最初からプライバシーを設計プロセスに組み込むための技術設計とエンジニアリングを促進するアプローチを指す。
- 仮名化（Pseudonymization）
  - この文書において仮名化とは、「データセットにおいて個々人が識別できるレベルで識別子の除去や置き換えをし、個人情報に修正をかけること」を指す。
- 再識別化（Re-Identification）
  - この文書において再識別化とは、「匿名化されたデータを再識別化により個人や団体と結びつけるプロセス」を指す。
  - UN Global Pulse Data Innovation Risk Assessment Tool guidance によれば、個人や団体の識別の不可を決定する際、個人や団体の識別が発生する可能性があるあらゆる場合を想定するべきだとされている。
  - 再識別化の可否に関する要素は、専門知識やコスト、時間、テクノロジーなどがある。
- センシティブなデータ（Sensitive Data）
  - この文書においてセンシティブなデータとは、人種や民族、政治的意見などに関連するデータを指す。

付録 A：どのようにしてデータ分析は SDGs を支えることができるか

① 貧困をなくす
② 飢餓をなくす
③ 健康・幸福
④ 質の高い教育
⑤ 性的差別撤廃
⑥ 綺麗な水と浄水
⑦ 安くて環境に優しいエネルギー
⑧ 経済成長
⑨ 産業、革新、インフラストラクチャー
⑩ 不公平への対処
⑪ 持続可能なまち・コミュニティ
⑫ 責任ある消費と生産
⑬ 気候に対するアクション
⑭ 海中生物
⑮ 地上生物
⑯ 平和、正義、強力な制度
⑰ 目標に向けたパートナーシップ

文責：三國 陸真