

第 3 回 共通基盤ワーキンググループ会合 議事概要

1 日 時 平成29年 7 月 6 日（木） 10:00～12:00

2 場 所 総務省第 2 庁舎 6 階 特別会議室

3 出席者

【委員】

北村 行伸（座長）、川崎 茂、永瀬 伸子、野呂 順一

【審議協力者】

美添 泰人（青山学院大学 経営学部 招聘教授）

山下 智志（統計数理研究所 リスク解析戦略研究センター長）

伊藤 伸介（中央大学 経済学部 教授）

内閣府、総務省統計局、財務省、文部科学省、厚生労働省、農林水産省、経済産業省、国土交通省、日本銀行、東京都、岡山県

【事務局（総務省）】

横山大臣官房審議官

統計委員会担当室：山澤室長、永島次長、上田次長、佐々木室長補佐

政策統括官（統計基準担当）：澤村統計審査官、宮内企画官、辻元統計専門職

4 議 事

（1）調査票情報等の二次的利用について

（2）その他

5 議事概要

（1）調査票情報等の二次的利用について

事務局から資料 1-1、資料 1-2、総務省政策統括官室から資料 1-3、総務省統計局から資料 1-4 に基づいて説明があった。その後、伊藤審議協力者から資料 2、山下審議協力者から資料 3 に基づいて説明があった。

「次期基本計画における基本的な考え方」については、概ね適当とされたが、意見等を踏まえて一部修正することとなった。

主な質疑は、以下のとおり。

《英国など海外の状況》

○米国では、同じ調査のデータであってもリモートアクセスが可能なデータとして開示される情報とダウンロード可能なデータとして開示される情報とは異なっている。つまり、セキュリティによって利用形態が分かれている。英国ではいかがか。

→（伊藤審議協力者）英国のパブリックユースファイルについては、教育用のデータしか提供されていないが、匿名化手法が適用されたライセンス型のマイクロデータの

場合、利用者がライセンスを取得すれば、データをダウンロードして利用することが可能である。提供されているライセンス型のデータの中には、小地域分析用の人口センサスのマイクロデータも含まれており、例えば、大学院生であっても、ライセンスを取得していれば利用可能である。

○学部学生や院生が幅広いデータに自由にアクセスして利用できることが必要だと考えている。日本の場合は、利用申請の入口のハードルが高くて利用範囲が広がらない状態ではないか。英国ではいかがか？

→（伊藤審議協力者）学生は、匿名化された様々なマイクロデータを利用することができる。また、大学院生の場合、指導教員も研究に関わるという形で、個票データにもアクセスすることができる。なお、個票データを用いた研究の成果を公表する前には、提供者側による分析結果のチェックが求められる。

○英国では様々な利用形態があるようだが、それぞれの利用形態の頻度はどういう状況か。利用しやすい形態は、何か。

→（伊藤審議協力者）ライセンス型のマイクロデータについてはダウンロードによって簡単に入手することができる。一方、個票データのアクセスに関しては利用申請の手続きが必要である。

○英国のオンサイト利用の個票では、ある程度の住所情報まで提供されているのか。

→（伊藤審議協力者）オンサイト施設で利用可能な個票データに関しては、住所といった個体識別情報は含まれていない。詳細な地域区分やセンシティブな情報を含んだ個票データについては、オンサイト施設やリモートアクセスでなければ利用することができない。

《公的マイクロデータ研究コンソーシアム関連》

○データ提供する際に、欠測値を補完するなどの機能は、利用者側のストレス軽減になるので、是非進めていただきたいと思うが、その場合、データはクリーンな形に補完されて提供されるのか、それとも提供されたプログラムを使って利用者がデータを自分で処理するのか。

→（山下審議協力者）両方のパターンがあり得ると思う。初心者用には処理済みのデータ、高度な利用者向けには、データクレンジング手法をメニュー化しておき、自分の研究に最適な手法を選択できるような形があるのではないかと考えている。

○利用者が使用するソフトウェアもいくつかの種類があるが、そこまで対応したデータベースになるのか。

→（山下審議協力者）ソフトウェアの組合せをどこまで増やせるのかは、課題と考えている。

《調査票情報の提供》

○調査票情報の提供・活用状況を見ると、個票を利用する形態の統計法第33条第2号による利用が急速に伸びているが、研究室におけるデータ管理の面で安全性が確保できているのかと危惧している。英国の状況はどのようになっているか。

→（伊藤審議協力者）英国に限ったことではないが、個票データはサーバ上にあるだ

けで研究室のパソコン上にはないというのが一般的であり、資格取得者しか個票データにアクセスできないという制約を課しているのが海外の状況だと言える。

- 旧法の目的外申請は、官学の共同研究の側面もあったと思う。現在の法第33条に基づき調査票情報の利用に当たっては、自分の場合、研究室のパソコンで個票を処理しており、データ管理には十分に留意しているが、匿名性確保の観点からは怖い面がある。条件の整備が必要であるが、法第33条第2号による利用をオンサイト利用に移行していくことが望ましい。

《オーダーメイド集計》

- オーダーメイド集計は集計表のみを提供するので、匿名性に関しては十分に安全かと言えば、そうとも限らない。実際、複数の集計表を組み合わせることで匿名性が失われる可能性もある。歯止めは、研究者の良心的な行動を保證する罰則規定の存在だと考える。
- 民間がオーダーメイド集計を依頼しようとする、何ができて何ができないのかが分からず、利用するまでに手間暇がかかる。イージーオーダー的なことができないだろうか。また、申請後利用可能になるまで何カ月もかかるのは利用しにくい。利用料金が実費であるため、事前に見積もりができないのもネックである。利用条件の緩和は大事だが、利用申請の分かりやすさ・透明性の観点も重要である。
- （統計センター）利用申請者からの問い合わせが多い事項などは、Q & Aにして提供するなど利便性向上に努めている。
- 民間事業者から要請の多い集計表は、e - S t a tに整備するなどの対応も必要ではないのか。

《匿名データ》

- 我が国の匿名データは、トップコーディングなどの匿名化手法を細かく公表しているが、データの真実性に重きを置き過ぎているのではないか。海外では、安全性と有用性を確保しながら匿名化手法の細かい部分は公表していない。同様な手法を導入しない限り、安全な匿名データの作成には多大な労力が発生する一方で、提供できるマイクロデータが限定される。この点に関する抜本的な改善がマイクロデータ活用のために必要である。
- なぜ日本の匿名データの利用は伸びないのか。労働データを分析する場合、法第33条に基づき提供される調査票情報は歪みもなく非常にいいデータである。一方、匿名データは、労働の分析に不可欠な地域情報がないなど使い勝手が悪い。安心して手続きも簡単に利用できるが歪みのない情報の提供は教育・研究上重要である。例えば、学生が簡単に利用できるが地域情報がないデータ、院生が申請で利用できる地域情報が付与されたデータ、限定された者が限定されたサイトで利用できる個票そのものなど、利用者の研究上の要請のレベルによって利用できるデータを分け、秘匿性を高めたデータは、より使いやすい形で提供することが必要ではないか。
- 日本の匿名データは、1種類しか提供されていないが、今後、利用者のレベルに応じて使えるように複数のデータを整備することを考えてほしい。

○分析の有用性を保証しつつ、秘匿性も維持できるマイクロデータを提供する仕組みを考えてほしい。そうすることで学生なども使えることになる。また、誤解が生じないように、秘匿に関する専門的な用語も整理が必要ではないか。マイクロデータの利用という意味では、オンサイトによる行政記録情報の利用も議論を進めてほしい。

なお、美添審議協力者から、審議時間が限られていることから、追加的な意見を書面で提出する旨、発言があり、以下のご意見をいただいた。

○オーダーメイド集計については、①具体例を提示して、利用者に費用対効果を予想させる、②要望の多い集計表はウェブ上で公表する、③利用し易いデータベースを作成し、統計作成者の負担を軽減させ、結果提供までの期間を短縮する、④メタデータの内容について、統計間で基準を明確にすることが必要である。

○教育用・一般用の Public Use Micro Data を作成する場合の視点は「分析の有効性と秘匿性の担保」の実現である。スワッピングなどの適用によって秘匿性を高めたデータセットについては、典型的な分析を実施した場合に、秘匿前のデータによる分析と類似の結果が得られることを確認すればよい。統計局・一橋大学が提供してきた疑似マイクロデータもこの候補に含めてもよいと考える。また、ウェブで利用可能とすることが望ましく、学生、企業などによる利用拡大の実現を目指すべきである。

○プログラム送付型リモートアクセスについて、プログラム点検にある程度の負担が発生することは避けられないが、英国のようにリスク評価の基準を示して、最終的な責任は分析者が負うことを明確にする必要がある。その目的は「データ提供者である調査客体に安全性を保証する」ことであり、この点からも、ある程度の罰則が設けられていることが望ましい。

○二次的利用に関するワンストップサービスについては、研究目的と委託集計が主要な対象となると考えるが、教育用・一般用であっても分析手法の指導など、適当な機会を設けることが望ましい。

(2) その他

次回の会合は、7月20日（木）13時から中央合同庁舎第4号館にて開催予定。

以上

<文責 総務省統計委員会担当室 速報のため事後修正の可能性あり>