

「オンサイト利用」及び「プログラム送付型の集計・分析」の 実用化に向けた課題整理と今後の進め方

平成 25 年 6 月 25 日

総務省政策統括官（統計基準担当）

1 背景

前回（平成 25 年 3 月 26 日）の研究会において、プログラム送付集計は発展的・応用的な課題としていたが、オンサイト利用施設への物理的なアクセスが難しい地方在住の利用者の利便性確保の手段として、プログラム送付型の集計・分析は有力な解決策の一つであると考えられる。

そこで、当研究会において、オンサイト利用に加え、プログラム送付型の集計・分析の実用化に向けた課題整理も行うこととしたい。

2 課題整理

（1）オンサイト利用

前回までの議論において、情報漏えい対策の観点から、現行の調査票情報の利用形態からオンサイト利用施設での利用に徐々にシフトし、オンサイト利用施設内では探索型の利用ができるようにし、成果物を持ち出す際に個人情報漏れることがないかを厳格にチェックするという方向性について、おおむね共通認識が得られたものと考えられる。この方向性に進めていく上での具体的な課題について、再度整理した。

○ 運用形態（「オンサイト利用に関する提案」平成 24 年 12 月 19 日 資料 1－2 から引用）

→ ①直轄方式：政府統計機関またはその委託を受けた独立行政法人（以下「政府統計機関等」という。）が、その建物内に設けたオンサイト施設の運用を行う方式。オンサイト施設設置機関は政府統計機関等である。

②委託方式：政府統計機関等からオンサイト利用の運用業務の委託を受けた大学法人または大学共同利用機関法人（以下、「運用受託機関」）が、その建物内に設けたオンサイト施設の運用を行う方式。オンサイト施設設置機関は運用受託機関である。

※ 業務委託は法人間の契約

・運用受託機関内でオンサイト施設の運用を実際に担当する部局・部署を明確にする。

※ 契約に当たっては、

・施設が技術的基準を満たさなくなったり、運用面で不都合が生じたりした場合は改善がみられるまで業務停止。

・組織が調査票情報等の不適切使用などを看過したりした場合は認定取消を条件とする。

○ 施設利用料

→ アメリカセンサス局では、利用料を徴収しているものの運営経費を賄うには至らず、保健統計センターでも運営費全体の5～10%程度しか賄えず、ハードウェアの購入に充てられる程度であり、興味本位での利用をけん制する機能としての意味合いが大きい。

カナダでは、大学がオンサイト利用施設を運営しており、利用料を取るかどうかは大学に委ねられている。

諸外国の例を参考としつつ、これまでの経緯等を踏まえ、利用料設定の考え方について、引き続き検討を進める。

○ 施設内で作成した分析用プログラムの共有

→ 公的統計のデータは、国民の共有財産として、幅広く用いられるべきものであることから、オンサイト利用施設内で作成した分析用プログラムは共有する方向で検討。

○ 利用機会の公平性の確保

→ オンサイト利用の推進と併せて、プログラム送付型の集計・分析サービスを提供することにより、地方在住の利用者のニーズを満たすこととしてはどうか。

(2) プログラム送付型の集計・分析

諸外国における実用例を参考に、我が国においても統計法第33条（調査票情報の提供）の一形態とし、プログラム送付型の集計・分析の導入の検討が必要と考えられるが、運用面では以下の課題がある。

○ テストデータの提供

→ 利用者側は、実際のデータに近い分布を保持し、データレイアウトも実際のデータと同一なテストデータの利用を求めるものと考えられ、また、プログラム送付型集計・分析サービスの提供者側としても、きちんと実行されるプログラムを受領することが望ましい。

→ 多くの統計調査で早期にプログラム送付型の集計・分析を利用可能とするため、架空のダミーデータを新たに作成するのではなく、調査票情報からリサンプリング等の必要な措置（ただし、項目の削除や符号の統合は行わない。）を施したテストデータを作成し、必要なセキュリティ対策を措置された環境下において、それを利用する方法が考えられる。

→ ドイツ連邦雇用庁では、オンサイト利用が原則であり、滞在期間が終了しても集計・分析が終わらなかった場合にプログラム送付型の集計・分析を認める形となっている。すなわち、オンサイト利用施設で研究者が実際にマイクロデータを利用することが、プログラム作成の訓練となっている。

なお、(独)統計センターが試行提供を行っている擬似マイクロデータについては、研究開発に相当程度の期間が必要であり、また、既に作成・提供済みの全国消費実態

調査の擬似マイクロデータについても、擬似マイクロデータとして提供されている項目が限定されており、元のデータと同一のデータレイアウトになっておらず、プログラム送付型の集計・分析用のテストデータに用いることは困難と考えられる。

○ 集計・分析結果の提供までに要する期間

→ 調査実施者側は、結果の秘匿性が確保されているかをチェックするため、結果提供までに相応の時間を要する。一方、利用者としては、回帰分析等の統計分析において変数の入れ替えを行う場合など、様々な試行錯誤を行いたいため、結果提供までの期間はできる限り短い方が望ましい。

→ 多数の集計表が作成された場合、秘匿性のチェックに時間がかかると考えられ、できる限り短期間で結果提供を可能とする方法（例：利用者が作成するプログラムに予め必要な秘匿性チェックのスキームを組み込むなど）を検討する必要がある。また、統計分析の結果のチェックには比較的時間を要しないと考えられるが、元のデータを復元できないようにする必要がある。

○ 利用可能なデータ

→ プログラム送付型の集計・分析の対象とするデータについては、調査対象の秘密保護等の観点やプログラムの作成・実行が容易かどうか等の観点から検討する必要がある。

(3) その他（利用者教育）

統計データの二次的利用の促進を図るため、大学等の高等教育機関におけるマイクロデータの取扱いに関する教育を推進する必要がある。その一環として、(独)統計センターで試行提供を行っている擬似マイクロデータについて統計調査の種類を追加する等、教育用マイクロデータの一層の開発・提供が必要ではないか。

3 今後の進め方

政策統括官（統計基準担当）は、本研究会において整理したオンサイト利用に必要な要件の具体化を進める。

また、(独)統計センターにおいて、プログラム送付型集計・分析の実用化に向けた技術的研究を進める。