

## 擬似マイクロデータについての論点（素案）

平成 24 年 12 月 19 日

総務省政策統括官室（統計基準担当）

1. 経緯

- （独）統計センターでは、公的統計の二次的利用の拡大を図るため、統計法第 33 条に基づく調査票情報の提供を受けて擬似マイクロデータ（教育用擬似マイクロデータ）を作成し、その試行的な提供を行っているところ。
- 統計法第 33 条では、「統計的研究」を行う場合に外部の者に調査票情報を提供することができる旨規定されている。擬似マイクロデータの作成及び試行的な提供は「統計的研究」として整理できるが、本格的な提供を行う場合、匿名データ等との関係整理を含め、その制度上の位置付けを改めて整理する必要。
- なお、本研究会の平成 23 年度報告書においても、「このような新たな技術的手法による擬似マイクロデータについて、その研究を進めるとともに、総務省においてその制度上の位置づけ等について検討を行うことが求められる」と記載。

2. 整理すべき点

（作成・提供の目的及び必要性）

- 何のために擬似マイクロデータを作成するのか。擬似的なものであるため、研究目的には適さないのではないか。教育目的（マイクロデータの取扱い・分析手法の習得等）であれば、個々のデータの分布や平均を調査票情報の分布や平均と同様なものにする必要はあまりないのではないか。（秘匿性を確保する観点からは、例えば、匿名データを基にした作成方法の方が技術的・業務コスト的に効率的となるのではないか。）
- 現在の匿名データは使いづらいつわられており、その理由として主に①セキュリティ要件が厳しいこと、②学生に利用させる場合人数分の手数料が必要、という 2 点が挙げられる。セキュリティ要件の緩和については慎重に検討する必要があるが、手数料についてはアカデミック・ディスカウントについて検討する余地もあると考えられる。  
匿名データがより活用されるための方策も検討するとともに、調査票情報、匿名データ及び擬似マイクロデータの関係（役割分担、活用方法）を整理する必要があるのではないか。

（法令面、調査対象の秘密の確保の観点）

- （独）統計センターの試行的提供においては、擬似マイクロデータ作成者が統計法第 33 条に基づく調査票情報の提供を受け、その研究成果として擬似マイクロデータを公表するという方法をとっている。この場合、擬似マイクロデータについて調査対象の匿名性が確保されている必要があるが、現行の（独）統計センターの作成手法で個人が特定されるおそれはないと言えるか。
- 今後、本格的な提供を行う場合、現在の試行的提供からどのような点が変わり得るのか。
- また、（独）統計センター以外の者が同様の取組（詳細集計表を基にした擬似マイクロデータの作成・提供）を行う場合、どのように考えるか。

(その他)

- 「擬似マイクロデータ」という名称は、秘密保護を前提とする公的統計の作成や、匿名データの提供等との関係で混乱を生じるおそれがあるのではないか。

## 【参考】

## 1. 「平成 23 年度 統計法施行状況に関する審議結果報告書」(抄)

(平成 24 年 9 月 内閣府 統計委員会)

## 【本編】

## II 各ワーキンググループの検討結果

## (III) 第 3 ワーキンググループ関係

## 3 審議結果

## (1) 重点的な審議課題

## ① 統計データの有効活用の推進

## i) 二次的利用関係

## (イ) 施策の進捗状況等に対する評価

○ また、法第 33 条第 2 号による調査票情報の利用については、調査対象の識別可能性や情報漏えいのリスクへの対応として厳格な運用が求められている中で、オンサイト利用や、匿名データ・教育用擬似マイクロデータとの関係整理も必要である。

## (ウ) 今後の施策の方向性等についての基本的考え方

○ また、二次的利用を取り巻く諸課題については、総務省の研究会における検討状況を注視していくこととする。なお、オンサイト利用や教育用擬似マイクロデータの検討に当たっては、コストやその負担、国民の理解や研究者に対する国民の信頼感にも密接に関係することに留意が必要である。

## 2. 「統計データの二次的利用促進に関する研究会 平成 23 年度報告書」(抄)

(平成 24 年 7 月 統計データの二次的利用促進に関する研究会)

## 7 二次的利用の推進に向けた取組の方向性

## (5) その他新たな技術的手法の検討

(前半略)

さらに、諸外国における政府統計データの活用の状況について幅広くとらえてみると、ニュージーランド統計局では、公表されている集計結果表から作成されたマイクロデータが利用制限なしに提供されていることがわかった。日本では、統計センターにおいて、同様の観点から作成する擬似マイクロデータ(※)の研究が行われている。

※ 擬似マイクロデータ：一般に、詳細化した集計結果を基に作成したデータ。個々のデータの分布や平均が、調査票情報の分布や平均と同様なものとなるように作成する。個々のデータは詳細集計結果を基に作成したものであるため、実際に存在する個人・法人等のデータではなく、これらのデータから個人・法人等が特定されるおそれは少ないと言われている。

一方、現行の匿名データについてみると、秘匿処理のために地域などの情報の欠損が生じていることに加え、調査結果が公表されてから一定の期間を経過したデータとなっている。

このような匿名データの欠点を補完し、匿名データと類似の利用が可能なものとしては、上記の擬似マイクロデータが有効であると考えられる。擬似マイクロデータは、データの分布や集計結果が調査票情報によるものと同様なものとなるように作成されるものであり、現在は教育における利用を想定したもので、小規模な統計調査では作成できない、研究目的には適さないなど不十分な点もあるが、利用目的に制限なく利用することができると考えられる。

このため、このような新たな技術的手法による擬似マイクロデータについて、その研究を進めるとともに、総務省においてその制度上の位置づけ等について検討を行うことが求められる。

また、擬似マイクロデータの検討と併せて、匿名データ等の統計データを提供する場合の暗号化技術についても、今後検討することが求められる。