

マイクロデータ利用に際しての「障壁」とは

公益財団法人 統計情報研究開発センター 坂部 裕美子

1 はじめに

公的統計のマイクロデータ利用制度(二次的利用制度)が開始されて数年が経つが、利用申請件数の伸びは近年鈍化しつつある。これにはいくつかの理由が考えられるが、マイクロデータ利用制度そのものの魅力が一種の「障壁」を超えて浸透しないために利用のすそ野が広がらない、ということもその一つと言える。この「障壁」は具体的にどのようなものか、そしてこれを取り払うにはどうすればよいかを、統計のライトユーザー(接した経験があるのは公表された統計調査報告書のみ、というユーザー)の視点を重視して考察し、対策を講じる。

2 マイクロデータ利用拡大を妨げている要因

① 新規顧客たりうる潜在的利用希望者への訴求が出来ていない

「ビッグデータ」という言葉が一般に定着し、従来の統計専門家以外で「統計を使った分析を試みたい」と考える研究者は、潜在的には増えていると思われる。しかし現在 Web 等で確認できる二次的利用制度についての説明は、用語が高度に専門的で、統計のコアユーザー(統計研究者や、統計局の関係者)でないと理解しがたいものになっている。そもそも「二次『的』利用」という言葉を使う理由さえライトユーザーには分からないであろう。

② 公開されている利用説明書が分かりにくい

新たな「利用希望者」が、マイクロデータ利用のメリットや制度の趣旨を理解し、これを使えば自分が想定するような集計ができそうだ、と確信を持った段階で、次に参照すべき資料は、申請に関しては「利用の手引」、個々の集計計画策定に関しては「データレイアウト及び符号表」である。いずれも Web 上で容易にアクセスはできるが、一読したのみでは、記載内容はほとんど理解できないと思われる(例えば「政府統計個票データレイアウト標準記法」の具体的な内容)。

3 今回の報告

幣財団では、自主研究会として「二次利用促進研究会」を開催し、新規利用者がマイクロデータ分析を行おうと考えたときに、その推進を阻害する「障壁」がどのような局面(例:利用相談、研究計画立案、プログラミング)で発現するかを具体策に検討し、研究会の成果物として「公的統計のマイクロデータ利用ガイド」を作成した。

今回は、この内容を一部引用しながら、二次利用の拡大に向けての対策を提案する。