

第21回統計データの二次利用の促進に関する研究会 議事概要

- 1 日時：平成28年1月15日（金）16:15～18:15
- 2 場所：総務省第二庁舎 6階特別会議室
- 3 出席者：廣松座長、縣委員、玄田委員、安田委員
田家政策統括官、吉牟田統計企画管理官、植松調査官
《オブザーバー》
内閣府（統計委員会担当室、経済社会総合研究所）、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、
独立行政法人統計センター
《審議協力者》
（独立行政法人統計センター）樁理事長
（統計研修所）小林教授
《事務局》
総務省政策統括官（統計基準担当）統計企画管理官室（佐藤補佐）
- 4 議題：（1）オーダーメイド集計の利用条件緩和等
（2）匿名データの年次追加に係る諮問手続見直し
（3）オンサイト利用の検討状況
（4）統計委員会の審議結果報告書に記述された今後の課題
- 5 議事の概要及び意見等
（1）議題1 オーダーメイド集計の利用条件緩和等
事務局から、資料1「オーダーメイド集計・匿名データ提供の見直しに係る省令等の改正内容とその利用者への周知について」の説明が行われた。説明後、意見交換及び質疑応答が行われた。

（意見交換及び質疑応答の概要）
○ （パブリックコメントの意見に対する考え方について）匿名化のプロセスの中で、例外的な人が1人いると特定化される問題があるが、それは仕方がないということか。
（縣委員）
→ 集計表の特定部分が1となるケースは、統計を公表する場合と同様、除外するのではないか。（安田委員）
→ その通りであり、個人の秘密の保護は保たれた仕組みである。（植松調査官）
→ それに類する記述を入れた方がいいのではないか。また、「個人」という表現も変えた方がよい。（安田委員）

- 統計学的には「調査対象者」という言葉を使うことが多いのではないか。(廣松座長)
- 利用者向け説明について、ユーザーの立場からみた関心は、実際にオーダーを出して、どのくらいのスピードで利用できるのか。その時に、資料に「金額は作業一時間当たり 5,900 円＋実費分」と書かれているが、クイックリスポンスに対してプレミアムを付けることが可能であれば、非常に合理的だし、大きなセールスポイントになると思う。(玄田委員)
- 1 回提供したデータの更なるケアというか、1 回 1 往復のリスポンスではなくてということもあり得る。(縣委員)
- 資料の金額は現行であり、同時に標準レートである。それ以外の対応に関してはケースに応じて考えるということはある。(廣松座長)
- 業務のキャパシティの問題で、やるのが望ましいけどできないということがあると思うが、プライバシーを確保しながら、信頼できる機関がクイックリスポンスの提供の代行やサポートを行ってフィーが提供される形があれば、手を挙げるアカデミックな機関は少なからずあると思う。(玄田委員)
- オーダーメイド集計が可能な統計調査は提供者側が決めるということだが、要望が出る可能性もあると思うので、それを受ける何かのルートも必要になると思う。(縣委員)
- 各府省がそれなりに対応しなければならないということで現状はこの程度に留まっているが、今後、二次的利用が盛んになり、かつ、各府省に対してニーズが多くなっていけば、提供する統計調査の種類や提供対象時期の拡大をしていかなければならないと思っている。(吉牟田統計企画管理官)
- このオーダーメイド集計の利用条件の緩和については、効果を発揮させるために、十分に民間等に周知することが重要。本日の指摘も踏まえ、十分な周知活動を行っていただきたい。(廣松座長)

(2) 議題2 匿名データの年次追加に係る諮問手続見直し

事務局から、資料2「匿名データの年次追加に係る審議手続簡素化に伴うガイドライン改正について」の説明が行われた。説明後、意見交換及び質疑応答が行われた。

(意見交換及び質疑応答の概要)

- 改正のポイントで3つの条件確認ができた場合ということだが、例えば次の対象と説明のあった就業構造基本調査で震災関連や介護・育児に関する事項が加わったような調査内容の変更については、「調査事項の変更が形式的であること」に抵触するため簡素化措置にならないことになるのか。(玄田委員)
- ここで言う「調査事項の形式的な変更」は極めて機械的なものを想定しており、それには該当しない。また、「母集団に変更がない」についても、5年に1度の調査だ

と、国勢調査の調査区が変わっていることが多いので、ここにも該当しない。したがって、①の「統計委員会への諮問を要さない」には相当しないと考えているが、例えば、統計委員会で実際に審議される際に、新様式により前回からの変更が本質的かどうか分かるようにして、議論が集約できるのではないかと考えている。(植松調査官)

○ 「匿名化手法の変更」について、トップコーディングとかボトムコーディングのしきい値の変更などは含まれるのか。(安田委員)

→ それは匿名化手法の変更と考えている。例えば、予め属性の5%を基準にトップコーディングするというような決め方であれば、その限りにおいて変更ないと言えると思うが、通常は何歳以上という形になっている。(植松調査官)

○ 「調査事項の変更が形式的」について、ガイドラインでは「(技術的な名称変更や選択肢の統合“等”）」となっているが、そこはルーティーン的に明確に判断が変わらないことを「形式的」という言葉で担保しているのか。(縣委員)

→ 形式的という言葉は、極めて機械的などという趣旨でとらえており、ここで例示として挙げている「統合」など以外にもあり得るかもしれないので「等」を使っている。(植松調査官)

○ 運用の段階では、個別の匿名データの作成段階で作成者と統括官室の方で具体的に詰める必要があると思うが、このガイドラインの改正については、本日の研究会終了後施行するという事なので、よろしくお願ひしたい。(廣松座長)

(3) 議題3 オンサイト利用の検討状況

事務局から、資料3「オンサイト利用の検討状況」の説明が行われた。説明後、意見交換及び質疑応答が行われた。

(意見交換及び質疑応答の概要)

○ 対象となる調査票の範囲はどこかで規定するのか。また、拡充する時には、どういう方針でどういう範囲までを想定しているか。(縣委員)

→ 1点目の調査票の範囲については、最初から全調査とはいかないため、予めオンサイト施設で使える調査票を明示することを想定している。2点目については、実際の執行の中で問題点を明らかにする面もあり、段階的に進めて行かざるを得ない。初年度はリソースの関係で数拠点という想定で、今まで統計データの利用の観点で密接に統計センターと協力いただいているところからまず声をかけ、段階的に声かけをしていき、ゆくゆくは手挙げ方式も想定して進めたい。(植松調査官)

○ 制度面で、対象範囲を何を根拠として決まり事を設定するのか。それを変更する時の決まりとか、オンサイトを拡大する際の最小限の機能・施設要件とかを決めないのか。(縣委員)

→ オンサイト利用は統計法第33条が根拠になるが、予め利用に供するものを示さなければならないわけではなく、あくまで調査実施者の判断に基づいて認めるというス

キーム。同様のスキームであるオーダーメイド・匿名データについては、分かりやすいように1年に1度ホームページなどでリストアップするようガイドラインで規律しているが、調査票情報の提供については、公益性が高いものしか認めていないので、そもそも広く知らしめるものかも論点となる。2点目のオンサイト施設の基準も、ガイドラインを示して規律を定めていこうと検討を進めており、引き続きこの研究会でも議論いただく。中身は、例えば施設が施設環境にあるとか、事故があった場合の対応など。(植松調査官)

○ オンサイト利用に関して平成28年度予算要求の状況を教えてほしい。(玄田委員)
→ 28年度においては、オンサイトを運用する人員の予算要求をして、認められたと聞いている。機械については、予算要求という形にはなっていないが、試行を行うだけの施設の目処はついていると聞いている。(植松調査官)

→ 試行運用環境については、統計センターの基盤システムの余剰リソースを用いた形で、本番環境を構築する検証等をしていくことになっている。29年度の予算要求に向けて検討を始めようという状況。(統計センター)

○ 例えば、大学がオンサイト施設を作るとして、その予算はどこが負担するのか。助成金のような形で出すのか、それとも施設の方で適切な予算措置をしてもらうのか。(玄田委員)

→ そこは論点の一つ。大学の施設なので、ある程度の負担を求めざるを得ない。だが、今回のリモートアクセスの仕組みでは、人を貼り付けなければならない個別型オンサイトと違い、もう少しリーズナブルなもので考えている。7月の研究会でも助成金という話も出たが、それも検討課題の一つ。コンソーシアムでも議論すべき話だと考えている。(植松調査官)

○ リモートアクセスという言葉が一人歩きしていると危惧している。ユーザーから見ると、自分の部屋にPCを持ち込んで、カメラがあれば使えるように思えるので、広報活動の際に注意していただきたい。(安田委員)

→ 統計委員会の法施行状況審議の場でもコメントしたが、オンサイト利用が軌道に乗った場合には、今までのように調査票情報の提供を受けて自分の研究室で使うという方式はなくなることを、利用者に認識してもらわないといけない。(廣松座長)

→ それゆえに、施設の条件を明確にする必要がある。(縣委員)

○ 諸外国では、オンサイト施設の中で利用できるデータの範囲やデータ操作の範囲と、リモートアクセスを通じて利用できるデータの範囲やデータ分析手法の範囲は違う。例えばデータリンケージが必要な利用はオンサイト施設の中でしかできないという国もある。できることの違いがあり得るという認識ができていないと思うので、実際に運用する時には、利用者に対しての注意喚起が必要。(小林教授)

→ そもそもリモートアクセスという方法が考えられたのは、オンサイト施設にデータを置くことを危惧して集中管理したいということからスタートしている。それを考えると、2種類のレベルを考えないといけない。一つは、今小林教授が言われたような

リモートアクセスの使い方と、もう一つは、オンサイト施設内でリモートにより全データという扱い方もできるような使い方。(安田委員)

→ 今回はリモートアクセスという言葉を使わない方がいい。オンサイト利用とさえあればいいのであって、広く認識されているリモートアクセスとは違う。(縣委員)

→ 確かにリモートアクセスとはいろんな意味でとられる表現なので、ご指摘の通り、あえてリモートアクセスと強調する必然性はないと思う。(植松調査官)

○ 利用者の権限をどう考えるのか。研究者の中には、例えば正月の三が日や土日に研究している方がたくさんいる。オンサイト施設の管理と利用を考えた時に、場合によっては利用者と施設の間に齟齬が起こる可能性がある。それがきっかけとなってオンサイトに対する否定的な意見にならないようにするためには、利用者にとって最適な方策をコンソーシアムで意見として汲み取る方策を考えないと、せっかく作っても研究の進展にプラスにならないことが危惧される。(玄田委員)

○ いま指摘があったように、検討すべき論点は山積みである。ただ、少なくとも公的統計マイクロデータ研究コンソーシアムが動き出すので、そこで出た意見をどう整理し、この研究会として受けとめるかという点も今後考えなければならない。当面は統計センターの協力で始まるが、政府全体の仕組みとして想定していることから、各府省もご協力いただきたい。本研究会においても制度面、技術面に係る課題を引き続き検討していく。(廣松座長)

(4) 議題4 統計委員会の審議結果報告書に記述された今後の課題

事務局から、資料4「二次的利用の手数料について」の説明が行われた。説明後、意見交換及び質疑応答が行われた。

(意見交換及び質疑応答の概要)

○ これは、なるべく早く対応すべき課題である。オンサイトについてはこれから始まるので、中長期で議論すべき点である。(廣松座長)

○ 手数料体系だが、オーダーメード集計だと、かかった時間について支払うということでアフターケアになっていると思うが、匿名データはそうならないので、アフターケアは想定されていないように感じる。匿名データの料金がファイル単位になっている根拠があるのか。(縣委員)

→ 法律上はいずれも実費を勘案した費用という定め方。匿名データについては、共通的事務経費は1,850円徴収しているが、一ファイル当たりの額は実際に作成にかかる実費分を頂くような整理で政令を定めている。(植松調査官)

→ より有効に統計が利用されることを目的と考えると、体系的にはどちらがいいのだろうか。変更を検討するつもりはないか。(縣委員)

→ その点も含めて論点集約の段階と考えている。法律事項の改正となると相当の準備が必要となるが、交通整理からした方がいいと考え論点を提示した。(植松調査官)

- オーダーメイド集計の場合は一人一人のニーズに応じて対応を変えなければならないので、かかった時間に応じて実費徴収する形だが、匿名データは一回ファイルを作成すると、後は誰が利用しても役所でデータを加工する作業が発生しないので、そういう利用形態の違いによって料金徴収の仕方が異なっている。その変更が今後出てくれば見直しが発生すると思うが、基本的にはこうなっている。(吉牟田管理官)
- 匿名データの内容によってファイルのコストが全然違うのではないか。現状はどんなファイルでも一律であり、内部補助をしているような形になっていて、ファイルごとの個別コストは考えてないように思えるが。(縣委員)
- まさしくそこが論点。実際匿名データを作るに当たり、加工度によってはコストが変わる可能性があるが、今経費負担を求める仕組みにはなっていない。(植松調査官)
- 匿名データの作成コストを受益者に負担させるのは難しいと考える。二次的利用のための供託金のような制度が必要なのではないか。調査実施者が匿名データ作成のコストを負担するとなると、拡大の制限とならざるを得ないし、一般財源でやる程パブリックな合意があるかどうか。先ほどのコンソーシアムなのか、データの潜在的・顕在的な利用者から費用を負担いただき、匿名データの作成と普及促進を図る方がいいのではないか。もっとも、コンソーシアムの普及・発展にマイナスかもしれないので、中期的に理解を得ていくべきもの。(玄田委員)
- このオーダーメイド集計や匿名データを始める前の検討段階では、そもそも統計調査とは税金で実施しているのであり、利用するのに更に手数料を取るのはいけりからんという意見もあった。とはいえ、オーダーメイド集計も匿名データもコストがかかっており、全て調査実施者に被せるのは、今の統計関係のリソースからいうと過大ではないか。利用者側の情報を使うことに対する意識の変化も必要。過去の統計の日の標語で「水と空気と統計と」というものがあつたが、これらは一般に無料だと思われているところがあり、情報に関する意識を変えていく必要がある。(廣松座長)
- 水や空気と同等なのは一般に公表されている統計であり、オーダーメイド集計や匿名データは違うのではないか。(縣委員)
- その通りだが、一般の反応では、オーダーメイド集計も匿名データも、元は同じと思われているのではないか。(廣松座長)
- オンサイトの普及と匿名データの提供で、どういうふうにユーザーが分かれていくのか将来的な読みが必要。将来的には、匿名データの主要なユーザーは高等教育目的、オーダーメイド集計は民間を巻き込んだ研究と分かれていき、いわゆる研究者のユーザーはマイクロデータを直接探索的に使えるオンサイト利用に移行していくという予想もある。それについて、課金の体制を議論する余地がある。(椿理事長)
- そのような形に誘導していく考え方もあり得る。(廣松座長)
- 私の一橋大学での匿名データ提供業務経験で言うと、統計データのレイアウトや調査そのものの知識を深く持っていない方からの利用相談で、そんな使い方はできないと思いつながら相談を受けていたことがあつた。オーダーメイド集計に関しても、窓口

担当者から聞くと、統計表の構造をよく知らない人もいるとのこと。申請の前にそういう知識を身に付けられる講習会を開くことも考える必要があるのではないか。それを受講した上で利用相談してもらえれば、結果を得るまでの時間を短縮でき、申出者と事務担当者お互いの負担を軽減できる。(小林教授)

→ 学会等の協力を得ないことには、政府だけでできる話でもないので、コンソーシアム等の場でも詰めていただきたい。(廣松座長)

○ 手数料に関する短期的な問題として、特に大学関係者からは高いという声強いが、それにどう対応すべきかどうか。(廣松座長)

→ 匿名データ制度創設初期の頃、一調査での分析が多かったが、最近は年次を並べる、他調査とのリンクなど複数のファイルを一人でオーダーされるところが、高いという印象になっていると思う。制度上可能かどうか分からないが、ボリュームディスカウントというアイデアもある。(安田委員)

→ 科研費を取っている場合には決して高い金額ではないが、皆が科研費を取っているわけでもない。取れるところからはしっかり頂いた上で、減免措置を十分に行う方が現実的と思う。(玄田委員)

○ 新統計法にオーダーメイド集計や匿名データの提供の項を加える時に、手数料の議論はあまり詰めてやっていない。これまで運用してきた意見をどういう形で今後反映するかというのが、この研究会の一つの大きな任務だと思う。本日、ボリュームディスカウントや、クイックリスポンス・プレミアムといったアイデアを頂いた。それらを現行の法律の枠の下で制度化できるかどうかについて直ぐに結論を出すのは難しいかもしれないが、本日の議論も踏まえて総務省政策統括官室において検討を進めてほしい。(廣松座長)

(5) その他

○ 3月に1回目のコンソーシアムの会議が開かれるということだが、参加される会員を紹介してほしい。また、28年度から試行されるオンサイト施設が具体的にどこになるのか教えてほしい。(厚生労働省)

→ コンソーシアムの出席者・参加大学は、今まで調査票情報を利用していた方や機関に対して、コンソーシアムの事務局から2月上旬以降、メールあるいは他の手段で声かけしていく。試行時のオンサイト施設については、今後整理するので、まだ明言できる状況ではない。(植松調査官)

(以上)

《文責：統計企画管理官付高度利用担当》