

第1回利活用検討会議事概要（主な意見）

1. 座長について

座長は、構成員の互選により、阿部 正浩 中央大学経済学部教授に決定。

座長代理は、座長の指名により、野口 晴子 早稲田大学政治経済学術院教授に決定。

2. 資料2「厚生労働統計等に係るデータ利活用の現状と課題」に関して

(1) オンサイト施設は令和元（2019）年5月より本格稼働し、まだ利用実績が少ない状況である。オンサイト施設の取組を拡充させていく上での課題はどのように考えるか。

(阿部座長) 今般の新型コロナウイルス感染症に伴う、オンサイト施設利用の困難化や、テレワークの促進とも関連し、在宅などからのオンサイト施設の遠隔利用も検討が必要。

以上とは別に、オンサイト施設利用を拡充する場合、設置地域の拡充、費用負担及び利用時間の制約などによる利用者側のコストが検討課題。

(野口座長代理) オンサイト施設の利用には反対。研究者が自身の研究施設からデータにアクセスする環境は非常に重要であり、オンサイト施設までの物理的移動は非効率。現在はオンサイト施設数が少なすぎて、データに対するアクセスが研究施設の場所に依存しているのは不公平。

(川口構成員) 以上の意見を踏まえると、今後の利便性向上とセキュリティーの両立を図るという観点から、政府部局内のサーバですべての計算処理を行い結果だけを取り出せる形のリモートアクセスシステムが望ましい。様々な制約からオンサイト利用が広がらず、結果として磁気媒体の提供も併存し、審査に時間がかかったり、実質的なセキュリティーが確保しにくいなど懸念がある。オンサイト利用拡大については政府全体の方向で決まっており、厚生労働省だけで解決できる問題ではないと思うが、検討会としても議論を進めていきたい。

一方で、オンサイト施設拡大を前提とするならば、過去の古いデータが利用できない統計調査については、古いデータについても早く利用できるようにしてほしい。また、統計法33条第1項第2号申請で利用できる統計調査や提供年次を拡大してほしい。

(川崎ワグサーバー) オンサイト施設は、まだ立ち上がって日が浅いことから、現状だけで利用実績の多寡を判断しきれない。

オンサイトを含む二次利用の広報周知の努力は理解したが、科研費の申請段階などで周知するなどの工夫が必要ではないか。また、厚労省ウェブサイ

トの改善効果が目に見えるほどになるには時間がかかる。引き続き情報提供に工夫いただきたい。

日本では、研究コミュニティにおけるマイクロデータ利用の経験が浅いことも、利用が少ない一因であると言われており、この問題はニワトリ（提供）とタマゴ（利用）の関係のように思う。厚生労働省におけるEBPMの推進などデータ分析のニーズの高い業務と並行して、オンサイトを含む二次利用の推進に息長く取り組んでいただく必要がある。

(2) 令和2（2020）年3月に、調査票情報等の利用に関して、ホームページの内容を改正したことによる意見。

(野口座長代理) オーダーメイド集計や匿名データが有料となっているが、調査票情報等に課金すべき。昨年訪れたスウェーデンでは個票データを使う研究者の所属機関に、データの整備・管理・運営・審査料を課金することにより、「研究施設から直接、統計省のサーバーにアクセスできるシステム」を構築している。研究倫理・罰則は厳しいが、当該システムで研究者は各種調査票情報のマッチングや、各種統計ソフトによる集計が可能で、統計省のチェック後に結果だけを手に入れることが可能となっている。なお、現在の日本のデータ提供状況で、ルールだけ厳罰化することには反対。

また、調査票情報等の利活用数が増加傾向にあるのは、素晴らしいが、一方でデータの審査に時間がかかりすぎている原因の一つに、審査に携わる担当者数や予算の不足があると思う。政府が科学的根拠に基づく政策立案・評価を本気で推進しようとするならば、そのための人的・金銭的資源を投入し、配分する決断を政府がする必要がある。

(川口構成員) ホームページの改善はとても良い取り組み。総務省が作っているmiripoもよくできているのでそちらへのリンクも適切な場所に張るとよいのではないか。

(川崎が「サーバー」) 調査票情報等の利用に関するサイトが、従前よりも上位レベルとなっており、利用者の探しやすさの観点から改善された。ページの作り方が、内容がほとんど変わらないにもかかわらず、やや細分化されている点は修正すべきではないか。

また、学会のホームページのリンク集にこのサイトのリンクを掲載するよう、学会に協力を求めることも効果的ではないか。

(3) 審査に関する現状と課題について意見

(阿部座長) 「確認・審査」の業務については、申請書類の不備があれば手間と時間がかかってしまうと思うので、事前にどういった点について確認・審査を行っているのか、申請者に伝えるべきではないか。

(野口座長代理) 政府の統計調査全体を省庁横断で見直すにあたっては、審査や

評価に関わる人的・金銭的予算をもっと増やすべき。そうでなければ、科学的エビデンスに基づく政策立案・評価は「絵にかいた餅」になりかねない。

(川口構成員) 現状を知るために、申請日から承諾日までの日数に関して平均、中央値を示すなど時間について報告してほしい。

「また、利用する調査事項（調査票情報）が、利用目的及び集計様式又は分析出力様式等から判断して、必要最小限となっており、不要と考えられるものが含まれていないこと。」という条件が、審査を長引かせる主因である。政府全体で見直すべきではないか。

(川崎ワグザバー) 提供済みの統計について業務を分析し、申請者に誤解や疑問の生じやすい書式などがあれば改善したり、FAQなどを掲載するなどして、申請段階からできるだけ不備のない申請を提出してもらえよう、引き続き効率化の工夫をしてほしい。

3. 資料3「アンケート用紙（案）」に関して

(1) 当省調査票情報の二次利用者の実態を把握し、今後の改善につなげることを目的としてアンケートを実施予定だが、意見をお願いしたい。

(阿部座長) アンケートの送付対象者だが、利用回数や在住地などを上手くコントロールして配布すべき。回答者の属性の違いによる意識の違いを見たい。

(川口構成員) アンケートの実施には賛成。全数調査にしたらどうか。仮にサンプル調査だとしても、サンプルサイズを大きくするべきで、仮にある問いに対する賛成が50%のとき、サンプルサイズ100で標準誤差は5%である。これくらいの精度がないと信頼できないのではないか。サンプルの抽出は無作為抽出としてその抽出方法を明確にしておくべき。

(川崎ワグザバー) 利用者の声を聞くことはよいことなので、ぜひ実施してほしい。調査対象数を、もう少し多くしたほうがよいのではないか。このアンケートは記述回答が重要であるので、できるだけ積極的に書いていただくよう依頼すべき。多様な意見を見て、状況を把握することは有意義。

(事務局) 御意見を踏まえ、統計法第33条については、アンケートの送付対象を100人程度としたい。アンケートの送付対象は、直近の状況を捉えたいこと、古いと連絡先が変わっている可能性があること等から、令和2年5月末から遡って対象とすることとさせていただけないか。また、基準となる期日又は期間、実施期間、アンケート用紙の提出期限については、とりまとめの遅延に応じ、後ろ倒しとさせていただきたい。

また、当初は20人程度であることを想定し、同じ所属組織の申請者を除外する等としていたが、サンプルサイズを100人程度に増やしたことに伴い、同じ所属組織の申請者は除外せず、地方公共団体を除き研究者を主とした「令和2年5月末までに利用承認され、提供した者（地方公共団体等の場合を除く）」について遡って選定。ただし、同じ申請者は除外。」とさせてい

ただきたい。
(構成員一同) 了

(2) 資料3 アンケート用紙(案)への意見

(川口構成員) 問1はそもそも周知されていることを知らない人もいるのではないかと。

(事務局) アンケート用紙に、調査票情報の二次利用等のホームページのリンクを明示する。

(川口構成員) 問8はマッチングキーとは何か知らない回答者も多いのではないかと。

(事務局) 研究者であれば、ご存じかと思い、原案では特に説明していなかったが、御意見を踏まえ、補足説明を付ける予定である。

(川崎が「バー」) オンサイト利用については、「使いたいか」どうか以上に、「知っているかどうか」、「分からないことはあるか」といったことも聞く必要があるのではないかと。

(事務局) 一般統計とならないよう留意しながら、アンケート用紙への反映をする。(アンケート用紙案の修正案を再提示)

(川崎が「バー」) この種の調査は定期的に行うのが望ましいが、改めて調査をするのは手間がかかるため、今後は利用者に書式を示して、申請手続き、もしくは二次利用終了時に同内容について記入・提出してもらおうとよいのではないかと。

4. その他意見

(野口座長代理) 公的な研究目的で政府統計情報を利用する研究者が、倫理上のルールを厳守すべきことは、国際的にも常識。日本では、利用させる側(行政側)に社会的責任が課されていて、データを貸し出すことへのディスインセンティブになってしまっている。データの利用拡大には、研究者に一定の理不尽にならない程度の罰則を科すことも仕方がないことかもしれない。

米国のCMS等でもclaim dataを貸し出す際、日本の統計法第33条、調査票情報の利用に相当する場合の審査には1年以上の時間をかけており、審査時間が長いこともある程度許容する必要がある。その場合、他の情報(国勢調査、税務情報等)との突合は行政が行い、匿名化されたIDで突合する等、省庁横断でのデータ構築の取り組みが望まれる。

今回のCOVID-19感染拡大の教訓として、データに基づく科学的エビデンスに基づく政策立案とその評価を省庁横断で推し進めて頂きたい。

(野口座長代理) 昨年毎勤問題で様々な課題が明らかになったが、政府統計を二次利用することのもう一つの意義は、研究者という「第三者」が解析することによって、サンプリングの問題点やデータ作成上の誤り等をチェックす

る役割を果たすことが出来る点にある。但し、統計調査によりデータを収集すること、そして、それを解析することの意味を、国民に理解していただく必要がある。この世の中に完璧なデータや、現状分析・将来予測も含め 100% 完全な研究なるものは存在しないからこそ、データ作成者と第三者である研究者が相互チェックしながら、出来るだけ質の良いデータ、そして、良質なデータに基づく正確で精緻な推定が行われるような環境を整備する必要がある。

データの作成・解析上、意図的でなく発生しうるミスに対し、あたかも作爲的になされたかのような批判ばかりを繰り返すだけでは、データの質も向上しないし、質の高い政策立案・評価のための科学的エビデンスを構築することも出来ない。長期的には、「統計」や「データ」に対する国民の理解を促すため、中・高等教育におけるカリキュラムを作っていくことも重要ではないか。

(川崎ワグザバー) 例えば、今後、雇用保険関係や職業紹介の個別データについても、オンサイト利用ができるようにしていく必要がある。官民データ活用基本法では、統計調査のマイクロデータだけではなく、行政記録のマイクロデータについても利用推進を求めているし、昨年毎勤問題に関連した統計委員会での議論を聞いていると、賃金政策、労働市場政策などでは、データ利用を推進し、それによってEBPMをしっかりと動かしていく必要がある。

せっかくオンサイト利用のように、秘密保護に配慮した利用環境が整備されつつあるので、行政記録の活用についても、このようなスキームに乗せることができるよう、厚生労働省を挙げて取り組んでほしい。