

第 22 回統計データの二次的利用促進に関する研究会 議事概要

- 1 日 時：平成 28 年 5 月 31 日（火）10:00～12:05
- 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、中村調査官
《オブザーバー》
総務省（統計局、統計委員会担当室）、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
（中央大学）伊藤准教授
《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（佐藤専門官、越補佐）
- 4 議 題：（1）オンサイト利用に係る検討経緯と今後の進め方
（2）リモートアクセス型オンサイト利用に係る統計法第 33 条の運用に関す
るガイドライン（試行運用版）（案）について
（3）オンサイト施設基準について
（4）持ち出し審査の観点について
（5）その他
- 5 議事の概要及び意見等
（1）議題 1 オンサイト利用に係る検討経緯と今後の進め方
事務局から、資料 1「オンサイト利用に係る検討経緯」及び資料 2「リモートアク
セス型オンサイト利用の試行運用に当たっての基本的な考え方（案）」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）
 - SINET で全部の大学が網羅されているわけではない。国立大学は大丈夫だが、特
に最近の私学は SINET に入っていないところもあるので、今後進めていく上で注意
事項としてほしい。（安田委員）
 - 以前の案だと、WEB でシンクライアントの仮想専用線だったと思うが、今の案は
専用線ということで、将来も変えない予定か、それとも仮想専用線にして設置場所を
増やしていくのか。（縣委員）→ 仮想 PC を設置してシンクライアントで接続する。物理的な回線として専用回線
ある SINET を使用し、技術的には V-LAN という方式で接続する。（統計センター）

- 最終的には、そうすればオンサイト施設を増やせるということか。(縣委員)
- 当面は今考えている方式で進め、試行運用を踏まえながら環境に合わせられるかどうか検討していきたい。(統計センター)
- SINET は専用線というイメージが違うので訂正した方がいい。外部から自由に入れる。また、VPN を通すということになると、中で流れる情報が監視対象外になり、大学側の管理者が関与できなくなることを嫌うところがあるので、今後進めるに当たって大学側と詰めてほしい。(安田委員)
- オンサイト施設内部の管理については、その施設が属する組織(例えば、大学など)の情報管理体制に従わざるを得ず、統計センターだけで完全に管理するということはできないという制約があるのではないか。(廣松座長)
- オンサイト施設の設置場所については、必要条件を決めて統一性を図るのではなかったか。(縣委員)
- 資料4でオンサイトの施設基準があるので、後の議題でその辺りの話をさせていただきたい。試行運用版なので、やっていく中で問題が出てくれば、本格運用までに見直す。(中村調査官)
- リモートアクセス型オンサイトという名称が気になる。前回の研究会でも、リモートアクセスとオンサイトの言葉の使い方が議論になったと記憶している。(廣松座長)
- できればリモートアクセスという言葉は消してほしい。ただ、(従前のオンサイトとの差別化の必要から)それもやりにくいので、オンサイトを前に持ってくるか、「オンサイト利用(リモートアクセスによる)」にした方が良いのではないか。オンサイト利用を前面に出した方が良い。(安田委員)
- それは私も賛成。むしろ、オンサイトはリモートではないのか。(縣委員)
- 既存のオンサイト施設は、サーバがそこにあるもので、リモートアクセスになっていない。趣旨としては、既存のものと違うということを表したかった。(中村調査官)
- 元々海外のオンサイト施設は、回線が十分でない時にできあがっているもので、その中で完全に閉じたところで処理するイメージで始まった。必ずしもリモートが前提ではない。今回リモートアクセスを表に出すと、それだけが一人歩きしかねないので、オンサイトの中のサブセットの一つというイメージにした方が良い。(安田委員)
- 提供する側と利用する側で受け止め方が違うように思う。利用者側は、自分の研究室から直接アクセスできるという受け止めかたするのではないか。今回の検討では、オンサイト施設を前提に議論している。どういう言葉が良いか、アイデアを頂いて詰めていきたい。(廣松座長)
- 既存型は今後継続するのか、リモートアクセス型に順次移行するのか。(玄田委員)
- リモートアクセス型のものを普及させていきたい。これが軌道に乗れば、最終的には既存のものはなくしていきたい。(中村調査官)
- リモートでしかアクセスできない方式が定着していればこの用語で良いのかもしれないが、当分の間は難しそうなので、研究室からできるというイメージを与えない方が良い。(安田委員)
- 私も公的統計マイクロデータ研究コンソーシアムのシンポジウムに参加したが、大変

熱気があり、ニーズが高まっていることを感じた。それをどうやって受け止めて実行に移すかについては本研究会が大きな責任を負っていると感じる。まずは試験的ではあっても動かしてみることが大事なので、セキュリティ面は確保しつつ、動かしながらブラッシュアップしていく考え方で進めてほしい。試行運用では統計局の調査のみ（特に全国消費実態調査）とのことだが、それ以外の統計調査に関しても大変要望が強い。本格的な運用に至った暁には、各府省のデータも含めて、より有益な研究が行えるようになればと期待している。（廣松座長）

（２）議題２ リモートアクセス型オンサイト利用に係る統計法第 33 条の運用に関するガイドライン（試行運用版）（案）について

事務局から、資料 3 「リモートアクセス型オンサイト利用に係る統計法第 33 条の運用に関するガイドライン（試行運用版）（案）」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）

- 定められたオンサイト施設基準とその他で補足的に決めていることを実行した場合に、どういう逸脱した行為が考えられるのか。また、施設ごとの差異が生じることはあまりないように思えるが、その可能性はどうか。（縣委員）
 - 極端な例で言うと、PC からハードディスクを持ち出すとか、持ち出し審査を意図的にすり抜けようとする事などが考えられる。施設基準に合致した施設を用意すれば、運用上の差異は生じないと想定している。（事務局）
 - 逆に言うと、施設基準を満たさないと施設たり得ないということ。（縣委員）
- 持ち出しという概念について、資料 1 の概念図によれば、その日に機械から取り出せず、後日データが送付される。それは「持ち出し」という概念ではない。むしろ「持ち出し」と言う方がリスクに聞こえる。この言葉はやめて、利用状況に合わせた言葉にした方が良い。（縣委員）
 - 持ち帰って検討する。（事務局）
- 不適切利用への対策は、提供禁止措置もいいが、それ以上に名前を公表するのが一番効く。それをオフィシャルに書いてもいいのではないか。不適切な行為として考えられるのは、例えば SINET だと大学の中の回線にグループを当てて抜き取ることが可能。それを防ぐ手段として VPN というのはあるが、そうすると各大学のファイアウォールや検知装置をフリーで通るので、それを嫌う大学がある。（安田委員）
- 資料 3 の事前審査のところで「特に・・・の条件を削除する。」との部分は、探索型の研究を可能にするために全項目提供するという事なのか。（小林教授）
 - 研究の内容を示していただければ、その範囲で読める調査事項は基本的には全部提供する。ただ、特段の必要がなければ ID 等の情報は提供されないで、マッチングに必要といった理由を提示されれば、それに応じて提供することになる。（事務局）
- 持ち出し審査のところの「中間生成物」は「個別の情報を含まないものをいう。」と別のガイドラインで定義されている。単に中間生成物というと、何でもありという

イメージを持ってしまうので、ガイドラインに明示しておくことが必要ではないか。また、中間生成物に個別情報が入っているかの判断は難しく、その審査が大変であれば、現実的な方法を試行の中で検討すれば良い。(小林教授)

→ 資料3における「中間生成物」の記述は、ここでは途中段階で持ち出すものとの意味合いで使用している。その審査については、今後検討していく。(事務局)

○ 利用期間終了後の処置について、ガイドラインでは1年後に廃棄とされているが、ねつ造問題などを考えた時に、プログラムやログファイルを責任を持って研究者本人や大学が管理しておくことが問われる。トラブル回避のため、ここは詰めた方が良い。(玄田委員)

→ プログラムやログファイルの扱いについては別途検討したい。(中村調査官)

○ 問題が発覚するのは時間が経ってからなので、証拠データを何年残すかは検討が必要。監視カメラの映像も残っていることを利用者に周知することも必要。(安田委員)

○ オンサイトは探索型の研究がしやすいのも特徴だが、匿名データになじまない企業データが分析できるのも趣旨。それに付随する項目をガイドラインに付け加える必要があるかもしれないので、注意しておいてほしい。(安田委員)

○ ご指摘のとおり、まだ詰める必要のある部分があると思うが、今回の資料3は試行運用版ということで1年くらいの期間のものと考えて頂きたい。試行の結果として、追加が必要となる部分、削除していい部分などの判断が可能になると思うので、改めて検討したい。議題の2は概ね案のとおりガイドラインを作成して差し支えないと判断する。頂いたご指摘は私の方で預かり、必要に応じて修正して政策統括官決定とさせていただきます。それを踏まえて試行に参加する大学との手続を加速してほしい。(廣松座長)

(3) 議題3 オンサイト施設基準について

統計センターから、資料4「調査票情報のリモートアクセス型オンサイト利用に係る試行時の施設基準(案)」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)

○ 施設管理者は機関長ではないと思われる。外部に委託したり、形式上の管理者となることを避けるためには、施設管理者の条件を定めた方が良い。(玄田委員)

→ オンサイト施設の管理者は、ガイドラインで「責任者や事務の体制が定められている」こととなっており、ガイドラインに基づいて調査実施者が定める内規に基づいて確認されることになる。(事務局)

○ 資料4の3⑦で禁止しているUSBメモリと、3③のリモートアクセス用USBとの関係はどうなっているか。(廣松座長)

→ ③のUSBは、シンクライアント方式であるため、PCとサーバを接続するためのプログラムが入っているもの。統計センターが作成するもので、オンサイト施設内の利用者PC1台に対して1つ。利用者が持ち歩くものではない。(統計センター)

- そのUSBは持ち出してはならないと決めてしまった方が良いのではないか。(玄田委員)
- 次の議題かもしれないが、オンサイトで研究する場合、膨大なログファイルができるので、それを研究室に持ち帰ってじっくり結果を分析し、次回のオンサイト利用に備えるという形になると思う。このログファイルも持ち出しになるのか。(玄田委員)
- Rなどを使った時に結果が出るログファイルは中間生成物の扱いなので、ログが発生する都度全部研究室に持ち帰るのは、難しい問題をはらんでいるが、今後試行において検討すべき課題。(事務局)
- ずっとオンサイト施設にいないと論文を書けないということになれば、大混乱になる。そこまで中間生成物としてみなすのかどうか。(玄田委員)
- 施設基準はオンサイト利用の取組を成功させるために必須であるが、施設側の負担が重くなりすぎると拡大の障害になる。本日の議論を踏まえて検討してほしい。(廣松座長)

(4) 議題4 持ち出し審査の観点について

伊藤准教授から資料5「政府統計における個票データの利用のあり方と分析結果のチェックについて」の、事務局から資料6「持ち出し審査のイメージ・内容について」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)

- 資料6の基準に従えば安全だと思うが、厳しすぎてクレームが付くと思う。審査側の手間を考えると、これに該当すればすんなり通せると思うので、これを標準審査事項として、それ以外を個別審査とする方法もある。(安田委員)
- 資料5にある侵入者モデルを前提とすると非常に重い審査になるので、ヒューマンモデルで審査するのが实际的。悪意を持つ者に対しては威嚇を行うが、善意の研究者を想定して審査を行う。その場合、資料5で説明のあった承認の制度や教育プログラムを徹底することが重要。公的統計マイクロデータ研究コンソーシアムには、この教育プログラムを作ってもらうことを期待している。悪意の研究者に対する威嚇として、カメラでの監視とか、プログラムの公開なども考えている。後者については、知的財産権の関係で交渉が必要だが、応じない研究者には審査を厳しくすることも考えられる。(椿理事長)
- ユーザの立場からすると、記述統計の平均と分散及び回帰分析の結果の公開ルールをまず詰めていくことが大事。サンプルサイズ、自由度、変数の作成に関するルール辺りに最低限の合意ができていて、他の細かい点は別途議論しないと試行に間に合わないのではないか。(玄田委員)
- イギリスのリモートアクセスは24時間利用可能とのことだが、日本の場合どういうことか。(縣委員)
- 日本の場合オンサイト施設に行く必要があるが、24時間開いているとは想定されない。分析の中間的な成果物を持ち出すとなると、いろいろ制約が出るので、技術的に

- 可能であれば、中間的な成果物だけは研究室でも閲覧できるようになると研究者としてもありがたいだろう。(伊藤准教授)
- イギリスで研究室からリモートアクセスが認められているのであれば、なぜ日本で想定するような問題が起きないのか。(縣委員)
- まず審査を受けた上で資格が取得できないと個票データにアクセスできないこと、二つ目に、事前に教育プログラムに基づく訓練を受けていること、更に不正な利用をすると罰則がかかることとイギリスの担当者は強調していた。(伊藤准教授)
- システムで制約するのではなく、運用上の規定で制約をかけて問題が起きないようにしているということか。(縣委員)
- イギリスは法律で資格が明記されていることが重要と考えており、その上で運用上の規制を設けている。また、イギリスを含め海外では、オンサイト施設でないとは利用できない個票データとリモートアクセスで利用できる個票データが分かれている。例えば、人口センサスの個票データはイギリスのバーチャル・マイクロデータ・ラボラトリ (VML) に行かなければアクセスできない。より粒度の細かい個票データはオンサイト施設でのみ利用できるとしている国もある。(伊藤准教授)
- リモートアクセスできる個票データは秘匿処理が行われているのか。(安田委員)
- イギリスの場合、秘匿処理は行われていない。(伊藤准教授)
- ESS のガイドラインは公表されている審査基準であると理解している。資料 5 の説明では、実際の審査結果は審査をする人の主観的な判断によるということか。(廣松座長)
- チェックする担当者のスキルが重要で、そのために講習会等を開催している。(伊藤准教授)
- 管理者は利用の仕方をパトロールしているのか。監視していることをアナウンスすることで不正利用を未然に防止できる。(玄田委員)
- 研究室からリモートアクセスするが、監視カメラなどは置いていない。研究者を信頼していることを前提に制度が成り立っている。(伊藤准教授)
- 研究者の知的財産権との関係で、パトロールすること自体が認められない可能性もあるかもしれない。(玄田委員)
- 事前審査を緩和する分、持ち出し時の審査をしっかり行う必要があるのは当然だが、利用面を考えると、この審査が使い勝手を大きく左右することになる。安全性と利便性のバランスを考慮に入れて、審査を行う必要がある。本日の議論を踏まえて検討を進めてほしい。(廣松座長)

(5) 議題 5 その他

事務局から参考資料 1 及び 2 の説明が行われた。

(以上)

〈文責：統計企画管理官付高度利用担当〉