

情報資産としての統計と政府統計データアーカイブ

森 博美*

要旨

平成19年5月に新統計法が法律第53号として公布された。同法は、旧統計報告調整法の諸規定もその中に取り込むことで、文字通りわが国における唯一の統計基本法規として位置づけられることになった。新統計法は、旧統計法によるこれまでの制度的枠組みを抜本的に見直し、今後のわが国の統計制度の展開方向を指し示すいくつかの新たな制度要素を含むものとなっている。新たに導入された制度要素に加え、公的統計（政府統計）をそれまでの行政利用のための情報から新たに「国民にとっての合理的な意思決定を行うための基盤」情報として捉え直し、この新たな政府統計観に立脚して制度構築を指向している点が、新統計法の大きな特徴である。

現在、新たにいわゆる「司令塔」として組織された統計委員会、委員会事務局、さらには制度官庁を中心に、導入された諸制度の詳細設計に向けた具体的取り組みが進行中である。本稿では、わが国の統計制度の理念的根拠としての新たな統計観の演繹的帰結として想定される一つの制度的仕組みである政府統計データアーカイブに焦点をあて、調査票情報の維持保管の現状における問題点、さらには今後のアーカイブの組織や機能のあり方といった制度設計に関して、諸外国の動向なども踏まえつつ考察した。

キーワード

アーカイブ、統計法、統計改革、二次利用、行政記録

はじめに

60年ぶりに統計法（昭和22年法律第18号以下、旧統計法）を全面改訂した統計法案は、5月16日に国会を通過し、23日に平成19年法律第53号として公布された。新統計法では、加工統計や業務統計を新たに規律の枠組みに加えた他、調査フレームの整備や行政記録の統計利用といった統計作成の基盤整備、さらには二次利用に関する諸規定の新設など、旧統計法に比べて法制度の枠組みを様々な次元で外延的に拡張した点がその大き

な特徴となっている [森(7)]。旧統計法が統計調査の企画・実施過程を中心にしたいわば統計調査法規としての性格を強く持っていた点を想起すれば、これらの法制度の枠組みの拡張は、今回制定された新統計法を旧法から質的に区別するものである。

今回の法改正の特筆すべき点として、政府統計観の本質的な転換がある。明治期以来わが国の政府統計は、第一義的に政府による行政遂行のための情報として作成、使用されてきた。これに対して新統計法では、「公的統計」（政府統計）を国民共通の情報基盤とするという新たな統計観に基づく制度構築が行われている。第1条の目的条項に謳われたこの統

* 法政大学経済学部

〒194-0298 東京都町田市相原4342

計観は、新統計法が新たに導入した諸制度に対して、その理念的根拠を与えるものである。

本稿では、まず、この新たな理念の演繹的帰結として想定される政府統計の情報資産としての位置づけを与える。次に、統計の情報資産としての意味について、主としてデータ論の観点から明らかにする。さいごに、これらの考察を踏まえ、その情報資産としての価値を制度的に保障するアーカイブの機能などシステム設計に係るいくつかの論点を指摘する。

1. 国民共通の情報基盤としての統計

旧統計法は、法の目的を「統計の真実性を確保し、統計調査の重複を除き、統計の体系を整備し、及び統計制度の改善発達を図る」（第1条）と規定していた。他の多くの行政法と異なり、そこでは、遂行されるべき統計行政行為の内容が法の目的として列記されていたに過ぎず、政府が何のために真実性の確保された統計を整備しなければならないかといった統計行政行為が目指すべき究極の達成目的を掲げたものではない〔森(4)22頁〕。

これに対して新統計法では、政府統計を政府、民間を問わず、社会のあらゆる構成員が「合理的な意思決定を行うための基盤となる重要な情報」と位置づけ、それが「国民経済の健全な発展及び国民生活の向上に寄与」すべきであることを法の目的として謳っている。そこでは、新たな統計理念に基づいて法が目指すべき目的とそれを実現するための統計の「体系的かつ効率的な整備」や統計の「有用性の確保」のための一連の措置の実行といった統計行政の任務とが明確に区別されている。同法がこのように法の目的とその達成手段とを明確に切り分けたことの意義は大きい。なぜならそれは、今回の改革で新設された一連の制度的枠組みの存立根拠であり、それらが実効性をもって機能するための今後のシステム設計、さらにはその運用の基本方針を規定

するものであるからである。それはまた、わが国の政府統計の進むべき方向づけを与えるものとして統計委員会において策定されることになる統計基本計画のコンテンツやその取り上げ方、さらには計画の実施状況に対する評価基準をも制約するものである。

2. 情報資産としての調査票情報の利活用

ところで、法案の国会への提案文書は、「体系的かつ効率的な整備」ならびに統計の「有用性の確保」に関して措置されるべき統計政策を、「社会経済情勢の変化に伴い、国民のニーズに柔軟に対応した公的統計の整備が要請されている状況等にかんがみ、公的統計の作成及び提供に関し基本となる事項を定めることにより、統計調査の対象者の秘密を保護しつつ、公的統計の体系的かつ効率的な整備を推進し、調査票情報の多様かつ高度な利用を可能とするための措置を講ずる必要がある。」〔総務省(6)47頁〕と記述している。ここに述べられている「調査票情報の多様かつ高度な利用」に関しては、利用の対象としての調査票情報の利用可能な形態での保管とその多様かつ高度な利用システムという2つの側面がある。このうち本稿では主として前者を検討課題とすることから、後者については、差し当たり以下の点を確認するにとどめる。

新統計法はその第3章「調査票情報等の利用及び提供」において、「調査票情報の二次利用」（第32条）、「調査票情報の提供」（第33条）、「委託による統計の作成」（第34条）、「匿名データの作成」（第35条）、「匿名データの提供」（第36条）、「事務の委託」（第37条）、「手数料」（第38条）と、調査票に記載された個体データの提供、利用に係る一連の条文を掲げている。

周知のように、旧統計法では、第15条の2を唯一の根拠に、調査票の目的外使用禁止規定の適用除外措置として調査票情報の利活用が行われてきた。今回の法改正で、政府統計

の二次利用に関して独立の章が設けられ、その具体的な制度化が行われたことは明らかな前進であると評価できよう。なお、このような形での二次利用の制度化には、1995年の統計審議会答申『政府統計の新中長期構想』並びにそれを受けた特定領域研究「統計情報活用のフロンティアの拡大—マイクロデータによる社会構造解析」（研究代表者松田芳郎—橋大学教授（当時））での海外における政府統計の二次利用の実態並びにその法制度に関する調査研究、さらにはそのフォローアップとしての匿名標本データの試行的提供による利用実績の積み重ねが大きく寄与しているものと思われる。

3. 調査票情報の組織的保管

社会に関する統計的認識のためのニーズは、利用主体によってもまた時代によっても異なる。過去のデータを今日的視点から改めて解析することで、問題発生の起源やその展開経過を統計的に検証することができる。これは、過去のデータが調査の企画時点では全く想定されていなかった新たな利用価値を持ちうることを示唆している。

ここで、わが国政府のこれまでのデータ政策を簡単に振り返っておこう。これまで、指定統計調査については、旧統計法第7条第二号に従い、予め承認を受けた集計結果の作成を「統計目的」として調査票情報が収集され、所定の集計作業完了後は、永年保管される結果原表（又は結果原表を転写したマイクロフィルム）以外の調査票等の調査資料については、主として秘密保護の観点から所定の保管年限終了後に速やかに処分されてきた。ここでは調査票情報は、基本的に所定の製表作業の完了とともに事実上その使命を終えるいわば「一過性」の情報として位置づけられてきた。

今回の法改正により、行政機関は、「国民が基幹統計に関する情報を常に容易に入手す

ることができるよう、当該情報の長期的かつ体系的な保存その他の適切な措置を講ずる」義務（第8条第3項）を課されることになった。なお、基幹統計以外の一般統計についても、第23条第2項が第8条第3項の準用を規定している。これらの条文は、統計調査等によって収集された調査票情報の取り扱いの点で、これまでの原則を本質的に軌道修正するものであり、その意味で上記の各項は、新たなデータポリシーへの転換の決意表明であり、それは同時に、第1条に謳われた法の目的の演繹的帰結でもある。

4. 調査票情報保管の現状

結果報告書の公表後も、保存された結果原表からだけでは対応しきれない多様な統計ニーズが存在するのは事実である。このような潜在的ニーズの存在を踏まえて統計審議会は、調査票の原票を転写した磁気テープ等について、ニーズに応じた保存期間の設定さらには利用に必要な関係書類の整備の必要性を提言している[統計基準部(2)67頁]。これは、調査票情報を基本的に一次利用のためのものとした一過性的情報としてではなく、永続的価値を持つ一種の情報資産として捉えたものである。

この答申を受けて各統計作成機関では、調査規則を改定し、永年保管の対象となる調査資料の範囲を、それまでの結果原表（又は結果原表を転写したマイクロフィルム）だけから、新たに「調査票の内容が転写されている電磁的記録」についても永年保管するように改めた。その意味では、今回新設された上記の2項は、こういった慣行を法律として追認したものであるともいえる。

わが国の統計機構は典型的な分散型として知られる。分散制は調査票情報等の調査資料の保管の面でも同様に貫徹しており、永年保管される磁気媒体の調査票情報については、保管データの形式、メンテナンス体制、さら

には関連するメタ情報の範囲等についても、全面的に調査機関の裁量に委ねられており、現時点では調査票情報の保管に関して、省庁間で文書合意された統一的なマニュアルは存在しない。

5. データアーカイブの必然性

データをただ保管するのではなく、それを時代を超えて有効性を持つ情報資産として位置づける場合、その保管（アーカイビング）のあり方も自ずと異なってくる。それは単なる過去の記録として保存されてさえいけばよいのではなく、何よりも利用可能な形態で維持管理される必要がある。そこでまず、保管される情報資産の内容さらには保管施設のあり方について考えてみよう。

第一に、アーカイブに保管されるべき情報として次のようなものが考えられる。まず、調査票情報の中には、欠損値あるいは誤記入といった不完全データを持つレコードも含まれる。そのため、一般に集計作業に先立ってデータの補定（imputation）を含めたエディティングが調査票情報に加えられる。このようなエディティング処理が何段階かにわたって行われる場合、そのいずれを保管データセットとするかについての統一的な指針が必要となろう。なお、この場合には、後に述べるような保管データの使用に耐えうる内容並びに形式面での要素を充足したものでなければならない。また、このような調査票本体情報の他にも、エディティングの履歴情報としての処理プログラムや符号（コード）表など一般にメタ情報と呼ばれる周辺情報についても系統的に整備、保存される必要がある。

第二に、アーカイブに永年保管されるデータは累積的に増加する。そのため、施設は可能な限り経済合理的な形で維持管理される必要がある。わが国における国家財政が逼迫の度を強める中、統計組織も定員削減により大きく弱体化している。そのような状況の中で、

個々の統計作成機関がそれぞれ独自にデータの永年保管体制を維持するのは、費用原則に照らしても必ずしも現実的とはいえない。統計作成について現在のような分散型を仮に今後も維持したとしても、以下のような理由から、収集された調査票情報の維持管理については、集中方式を採用するのが適当と考えられる。なぜなら、第一に、保管の責任主体を単一にすることで、保管される本体情報さらには周辺情報について、分散保管に比べてより統一的な整備を実現でき、第二に、調査票情報の集中管理により個体情報の漏洩等のリスクを最小化でき、第三に、データの維持管理に関して、集中化に伴う規模のメリットを生かすことができ、第四に、分散型保管の場合に比べて保管データに基づく新たな統計作成がより容易になると期待されるからである。

これらに関係したアーカイブの技術的業務としては、大量情報の保管のための装備面での対応、保管情報を常に可読状態に維持するための技術面での対応、さらには突発的な事故や災害などに伴うデータの喪失の防止のためのバックアップ方策等が考えられる。また、その周辺の業務としては、データの保管形態、メタ情報として整備すべき要件等についてのガイドラインの策定並びにその更新がある。また、後に触れるが、アーカイブの統計作成機能の前提となるマッチング情報の整備やその安全な管理などもデータ保管に係る業務に属すると考えられる。

一方、制度的業務としては、統計作成機関や行政情報保有機関からの情報の移管に係るシステムの構築、秘密保護担保のためのアーカイブにおけるデータハンドリング従事者等の守秘義務の制度化、データリンケージのためのマッチングキーの使用に係る法制度の構築などが準備されなければならない。

6. データの保管形態とデータベースの構造

保管された統計データが情報資産として十

全に機能しうるには、データは収集された調査票情報の情報喪失を最小限にとどめる形態で保管される必要がある。かつて「統計=集計表」という認識が支配的であった時代には、統計は結果原表という形で保管され、追加集計もそれが許容する範囲内で行われてきた。結果原表は、全調査事項を網羅的に含むものではなく、個体ベースで調査票情報を保管する場合に較べて、明らかな情報喪失がある。さらに、結果原表という形で集計された保管形態の場合、回帰分析等の非集計的利用には全く対応できない。このような理由から、調査票情報の個体ベースでの保管が、その多面的な利活用の保証という意味でより適正な保管形態であると考えられる。

ところで、現在は統計における秘密保護のため、調査票情報から識別子を除去した個体情報が永年保管されている。当該調査結果を単独で再使用する限り、このような情報形態でもそれなりに使用に耐えうる。しかし、それを他の調査の調査票情報あるいは行政個人情報との連携使用を考えた場合、本質的な情報ロスがある。なぜなら、リンクのためのキー情報の喪失により、個体情報を相互にリンクすることで得られる追加的情報獲得の途が塞がれているからである。なお、データベースとして保管された個体情報のこのような利用可能性については、すでに今から30年以上も前にAukrustとNordbottenが、識別可能な形態での個体データを保管することで源泉やその把握時点を異にするデータ相互間のリンクによって将来のデータニーズにも対応できる「統計ファイルの体系」[(1)p.201]として提案している。

個体データの場合、当然のことながら、集計データに比べて個体が識別されるリスクが大きい。他方で個体情報をマッチングにより相互にリンクするには、何らかのリンクキー情報が不可欠である。そこで、個体が識別されるリスクを可能な限り排除しつつ統計相互

のリンクの可能性を確保するには、例えば、調査票情報を識別子情報と調査票本体情報とに分離し両者に秘密鍵としてのマッチングキーを付与しそれぞれ独立したデータセットとしてデータベースに保管し、特に識別子情報と秘密鍵情報については、最高度のファイアウォールの中に管理するといった方法なども検討の対象となりうるであろう。その場合、秘密保護との両立を図るために、秘密鍵情報の使用についての極めて厳格な制度運用が求められることはいうまでもない。

同一統計系列の異時点間の縦断面リンク、異種の統計相互あるいは調査データと行政記録情報とのクロスセクショナルリンクについては、統計作成の公益的必要性や結果の公開等の原則並びに適切な手続を経てそれを使用するというシステムが考えられる。この場合、秘密保護さらには調査票情報の適正利用の観点からも、マッチングによる統計作成申請の有資格者（機関）の範囲については、法律等で明文化される必要がある。また、「申請—承認」の手続きについても、作成される統計の利用面での公益性を含めた申請手続きの第三者機関への提出、承認を義務づけ、第三者機関はその適否について公開審議の上、適当と認められた場合に、理由を付してその申請案件についての承認を公示するといったようなシステムが考えられる。なお、具体的なマッチング作業はアーカイブの一部門（統計作成部門）に所属する限定された職員の専従事項とし、その業務担当者には通常の統計従事者に課せられる守秘義務に加えて特に厳格な守秘義務を課す必要がある。

7. 調査票情報の二次利用とアーカイブの諸機能

(i) オンデマンドデータ処理サービス機能

政府統計の二次利用の一形態として、利用者からの申請による独自集計がある。これについて新統計法は、「委託による統計の作成

等]として、「業務の遂行に支障のない範囲内」との条件つきながら、行政機関あるいは独立行政法人等に対して、「学術研究の発展に資すると認める場合…一般からの委託に応じ、その行った統計調査に係る調査票情報を利用して、統計の作成等を行う」（第34条）ことができる旨を新たに規定している。なお、このような委託による統計作成は有料で行われ、それについては同じく第38条がデータ処理に係る実費徴収を定めている。なお同法は、手数料の料率を政令に委ねている。

有料であるとはいえ、このようなオンデマンド型のデータサービスは、現状のような分散型体制の下では、業務の遂行に支障をきたすことを理由に十分に機能せず、この規定が事実上死文化する虞もないとはいえない。それが実質的な意味で機能しうるためには、それを一括受託し、オンデマンド処理についての経験についても蓄積が可能なアーカイブの設置が現実的であると考えられる。

ところで、企業・事業所データについてはその匿名化は極めて困難であり、また個人や世帯データについても特に微妙（sensitive）な内容の変数を含むデータセットが将来アーカイブに保管される可能性もありうる。海外での事例が示しているように、この種のデータについては、外部利用者に処理プログラムの提供を求め、組織内部の職員がデータの処理を一括して行い、その出力結果を審査の上で利用者に提供するというオンデマンド型処理サービスが有効である。

このようなオンデマンド型データ処理サービスの制度化に際しての当面の検討課題としては、利用料金の設定の他にも、利用申請並びに処理結果の審査に係る組織、さらにはサービスの対象となるデータの範囲の特定などが考えられる。

(ii) ミクロデータの作成機能

個体データについては、データの利用価値と秘密保護とが相互にトレードオフの関係に

ある。このために、データの秘密保護を図りつつ集計表や結果原表などよりは情報量の多い個体データを有効に利用するために、海外では1960年代から個体レコードに匿名化措置を施した標本再抽出データ（ミクロデータ）が作成、提供されてきた。それは、既存の集計結果表のみに基づく分析からは得られない様々な新たな知見を世に提供してきた。

国内外でこれまで行われてきた個体データの提供、分析事例を見ると、一口に個体データといってもその匿名性のレベルに応じていくつかのタイプが考えられる。①記入済み調査票そのものあるいはそのマイクロフィルム版、②調査票の内容を磁気媒体に転写したものの、③②から氏名や住所等の識別子情報を削除したもの、④③から利用者の申請に係る変数以外のデータを削除したもの、⑤③に様々な方法や強度の匿名化措置を施したものがそれである。

海外での提供事例が示しているように、これらの個体データについては個体識別のリスクが異なることから、その提供システムも自ずとそれに対応したものとなっている〔松田他(3)〕。例えば英国のLS (longitudinal study) データや企業データのように識別の危険性が高いあるいは微妙な変数を含む個体データセットの場合、外部からの分析提案を受けて組織内（on-site）に研究プロジェクトを立ち上げ、データ処理を内部の担当者に限定することで秘密保護と両立する形で利用者の分析ニーズに応えるオンデマンド型利用システム等が採用されている。

識別情報は除去されているものの匿名化措置は施されていない個体データの使用システムとしては、臨時的に公務員の身分を付与し守秘義務を課すことで秘密の保護を図る宣誓職員制度が一般に知られている。上記⑤についても匿名化の程度に従っていくつかのタイプがあり、弱い匿名化措置を施し利用者に適正使用の誓約書等の提出を求めて提供される

ものやアメリカのセンサスマイクロデータ（PUMS）のように、強度の匿名化措置を施した上で、磁気媒体で一般に販売あるいはウェブ上で自由にダウンロードを認めるマイクロデータもある。なお、このようなマイクロデータの提供チャンネルとデータアーカイブの関係については、本稿末の付図を参照されたい。

新統計法の第35条並びに第36条は、行政機関あるいは独立行政法人に対して、自らが行った統計調査の調査票情報から匿名データ（マイクロデータ）を作成し、学術研究目的での利用に対して有料（第38条）によるその提供ができる旨を規定している。このように、新統計法はマイクロデータの作成、提供を各統計調査の実施機関に委ねている。しかし、わが国の統計作成機関の中にはこれまでマイクロデータの作成経験を全く持たないものも少なくない。このため、これらの規定の制定を契機にわが国でマイクロデータの作成、提供が一気に加速するとは考え難い。

厳格に運用されるオンデマンド型処理サービスによってはじめて利用可能となるデータから公開マイクロデータ（PUMS）まで一口に個体データといっても実に多様である。秘密保護の観点からもまたアーカイブが保管する各データセットに対して統一的な運用原則に基づき最適かつ有効なデータサービスを行うためにも、将来的には調査票情報を集中的に維持管理するアーカイブにマイクロデータの作成機能を持たせることが、政府統計の二次利用のための情報提供システムとしては好ましいように思われる。なぜなら、マイクロデータの作成機能をアーカイブに集中することで、統一的基準に従った匿名化措置をデータに加えることの方が、個々の機関毎に独自の方法による匿名化よりも、データの秘密保護に関してより適切であり、それはまた各調査機関の業務負担の軽減にもつながるものであるからである。

わが国ではこれまで調査票情報の目的外使

用制度による個体データの使用には厳しい公益性が要求されてきた。それは、結果的に、調査票情報へのアクセシビリティの面で著しい官民格差という形でのデータ・デバイドの一因ともなってきた。マイクロデータの作成、提供機能のアーカイブへの集中化は、匿名化という技術的観点からだけでなく、特にわが国では個体データの提供方式においてかつて見られた一種の不透明性を払拭し、政府統計の二次利用におけるデータへの公正なアクセシビリティの保証という点からもその意義は大きい。

マイクロデータの作成、提供機能と関連したアーカイブの業務としては、秘密保護並びにマイクロデータの利用可能性との関連での匿名化技術の研究、利用目的に応じたマイクロデータの作成等が考えられる。なお、学術目的でのマイクロデータの利用申請の受け付け、提供等の業務については、諸外国の事例が示すように、大学などのしかるべき機関に業務委託することが望ましいが、わが国でもこれについてはすでに一定の経験並びに技術の蓄積が行われている。

(iii) アーカイブの統計作成機能

戦後、わが国の政府統計の再建における重要な課題は、統計の体系的整備であった。しかし、ここでの統計の「体系」は、分野別あるいはセンサスと速報統計といった統計の形態別の網羅的整備を意味した。そこでは、標本調査は母集団概念を媒介にセンサスと集計量ベースで関連づけられてはいるものの、個々の統計はいずれも単体（stand alone）として、統計体系の一要素を構成してきたに過ぎない。そのような中、個々の調査企画も単体の調査として提案され、それに対して制度官庁が主として報告負担軽減の観点から審査業務を遂行するという形でこれまでの統計調整行政は行われてきた。そこには、個々の調査客体そのものを体系的に把握するという視点が完全に欠落していた。

予算についてもまた投入可能な人的資源についても右肩上がりの状態を享受できたかつての高度成長期とは異なり、今後は厳しい資源制約とますます深刻化する調査環境の中で、拡大、多様化する統計ニーズへの対応を日本の政府統計は求められるであろう。

それへの対処策の一つとして考えられるのが、情報資産としてアーカイブに保管された既存の調査票情報の有効活用による統計作成である。まず異種の調査データを個体ベースで相互にリンクすることで、追加的な調査を実施することなく必要な統計情報を確保することができる。また、同一の調査データを個体ベースで時間軸に沿ってリンクすることから縦断面データも作成できる。特に企業や事業所を対象とする多くの経済調査の場合、同一調査について保管されている個体データのリンケージによって作成できる縦断面データからは、横断面（クロスセクション）データあるいは時系列データからは読み取れなかった新たな構造的変化に関する知見が得られる。パネルデータと呼ばれる新たなタイプのデータ整備は、わが国の政府統計が諸外国に較べて最も立ち遅れているデータ領域であり、統計の体系的整備には欠くことのできない要素であるといえる。

分散型統計機構の下で調査票情報が作成機関に事実上帰属することから、省庁横断的なプロジェクトとしての異種統計間のリンケージはまだ部分的にとどまっている。なお、一部の府省で内部的に企業統計について同一調査を個体ベースで時系列的に接続することで縦断面型データが作成されてはいるが、このようにして新たに作成された統計についても、その存在さえ一般には知られていないケースが多い。その利用も現時点では基本的に組織内部での利用が中心であり、例えばオンデマンド型のデータ処理サービスという形でその利用システムが制度化されているわけではない。

もう一つの対処策は、この間海外諸国がいち早く取り組んできた行政記録の統計への活用である。行政記録の統計利用に関してここで特筆すべきは、それが単に業務統計としてこれまで断片的に提供されてきたものの範囲を単に拡大するにとどまるものではないということである。特に動態的な統計情報という性格の強い行政情報を静態的な属性情報を持つ調査データとリンクすることで、これまで得られなかった全く新たなタイプの統計データの経常的な確保が可能となる。これについては、わが国でも遅ればせながら、今回の改革で新統計法に関連規定（第29条）が盛り込まれることで、ようやくそれへの取り組みの足掛りが作られた。

ちなみに、海外では、個体ベースでの調査間のリンケージ、さらには調査データと行政情報とのリンケージにより、パネルデータも含め、新たなタイプの多様な系列が作成され、オンデマンド型のデータ処理サービス等の形態で研究目的での利用も含め広範に使用されている。

むすび

政府統計については、巨額の子算と人員さらには被調査者の報告負担によって収集された個々の調査票情報に基づき所定の集計結果表が作成、公表されてきた。これがいわゆる調査票情報の統計目的での使用、すなわち一次利用とされてきた。ところで、一次利用されている情報は、収集された調査票情報が持つ潜在的に利用可能な情報量の総体からすれば、その一部に他ならない。情報技術の発展は、調査票情報の一次利用以外の活用可能性を現実のものとし、1960年代以降、世界の多くの国で研究者を中心に二次利用によるその有効活用がはかられてきた。匿名化措置により調査票情報が有する情報量は部分的に減じられるものの、それをマイクロデータとして広範な利用に供することで既存の公表結果か

らだけでは得られなかった多くの新たな知見を二次利用は提供してきた。

このような二次利用の広がりには、単に当該調査の利用可能性を拡大するだけでなく、その遡及的利用ニーズ、異種の調査データ相互のリンケージによる新たな利用可能情報の創出といった形での展開を生み出すことになる。このような二次利用という形での調査票情報の有効活用は、当然のことながら、それにふさわしい政府のデータ政策を要請する。それを象徴するのが、調査票情報を一次利用のためだけの一過性の情報としてきたものを、時代を超越して利用可能な一種の情報資産として位置づけ直すという新たな視点がそれである。

識別子情報を削除した調査票本体情報のみの保管というわが国の現行方式は、統計に係る秘密保護を最優先するとのこれまでのデータポリシーを前提すれば妥当な措置といえないこともない。ただ、収集された調査票本体情報さらには今後行政情報の統計利用が本格的に制度化され、また行政个体情報のアーカイビングが実現した段階では、そこでの情報喪失は、それらが情報資産として潜在的に有する情報の「統計作成目的」での有効活用、すなわち、孤立した一過性の情報から永続的かつ統計相互の連携による新たな情報創出といった展開可能性に対してその道を事実上閉ざすことになる。

もちろん、たとえば「統計作成目的」とはい

え、保管された調査票情報のこのようなリンケージ使用については異論も少なくないであろう。被調査者による調査票情報の提供はこのようなリンケージによる統計作成までも許容したのではなくそのためのinformed consentが必要であるというのは正論である。これは、統計目的というこれまでのデータポリシーの基本的立脚点でもあった。

本稿で論じてきたアーカイブの諸機能については、すでに諸外国ではかなりの程度現実のものとなってきている。とはいえ、わが国で今後それらを実現していくためには、本稿でも紹介したような多くの課題を着実にクリアしていく必要がある。アーカイブをひとつの理念型のような形でここで敢えて問題提起したのは、それが新統計法の目的規定に謳われた新たな統計観並びにそれに立脚する「国民経済の健全な発展及び国民生活の向上」の演繹的帰結のひとつとして想定される制度要素であるからに他ならない。本稿で提起した政府統計データアーカイブの構築に向けての検討課題の中には比較的容易に対応が可能なものもあれば、わが国固有の統計的風土の中でその合意形成に気の遠くなるような時間と努力を必要とするものもあると思われる。21世紀の政府統計機構を展望するに、その中で中心的な機構としての政府統計データアーカイブの構築に向けて、本稿での指摘から漏れていると思われる懸案事項も含め、ひとつ一つ粘り強く着実な取り組みが求められる。

参考文献

- (1) Aukrust, O., and Nordbotten, S. (1973.6) Files of Individual Data and their Potential for Social Research, *the Review of Income and Wealth* 19-2.
- (2) 総務省統計局統計基準部 (1995) 『統計行政の新中・長期構想』
- (3) 松田芳郎・浜砂敬郎・森 博美編 (2000) 『講座ミクロ統計分析1—統計調査制度とミクロ統計の開示』日本評論社
- (4) 森 博美 (2005) 「「統計法」と法の目的」法政大学日本統計研究所『オケージョナルペーパー』No. 12
- (5) 吉田 央 (2007) 「韓国「統計法」改正」法政大学日本統計研究所『統計研究参考資料』No. 95
- (6) 総務省 (2007) 「平成19年第166回国会 統計法案関係資料」

- (7) 森 博美 (2007) 「新統計法の成立とわが国政府統計の今後の課題」『計画行政』日本計画行政学会 30-4

Statistical Reform and Data Archive

Hiromi MORI

Summary

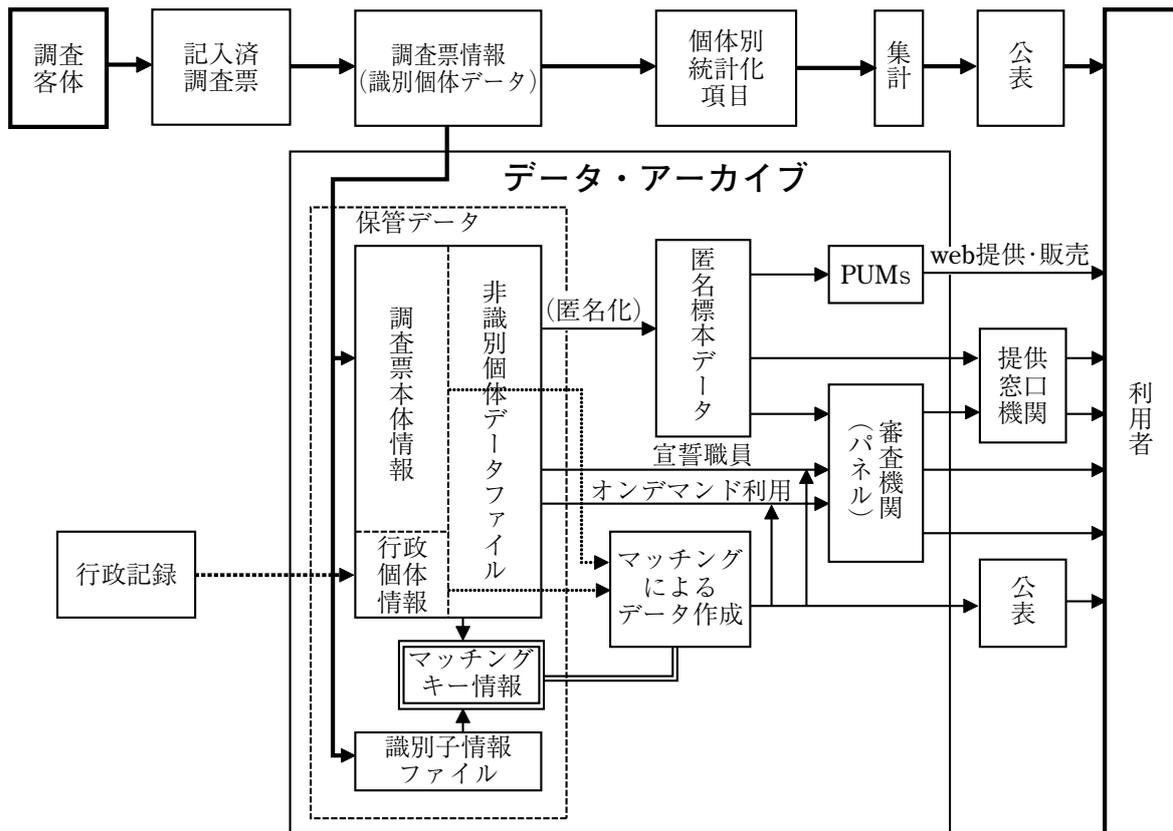
The new statistics law (Law No. 53 of 2007) in Japan was promulgated in May 2007 and will come into force in full scale on April 1st 2009. The law is characterized, among others, by expanding its scope in terms of application in several dimensions in which Japanese statistical system has been lagging behind based on the new concept that official statistical data should be a kind of public goods produced and disseminated not exclusively for governmental purposes as is the case at present but also for the general public as substantial information for their rational decision making.

Article No. 8 (3) of the new law requires those who represent Government bodies to take appropriate measures for the long-range and systematic storage of statistical information to endorse easy access by the public. One may reasonably assume its systematic archiving as the deductive institutional outcome of the new concept of official statistics.

Aim of this paper is twofold. First, to examine the current practice of storing statistical data by the Government and bring under light some problems inherent in the present system. Second, to bring forward an institutional design of the archive including its setup and its functioning. As the paper has elucidated, the Government statistical data archive is expected to play a substantial role not only in keeping data for historical analyses but also for multifarious secondary use and creating new statistical data through matching different set of cross-sectional survey records as well as surveys and administrative records.

Key Words

archive, statistical law, statistical reform, secondary use, administrative records



付図 統計の作成，提供とデータアーカイブの概念図