

「統計データの二次的利用」及び「統計データ・アーカイブ」について

平成 25 年 7 月 9 日

総務省 政策統括官（統計基準担当）

1. 「統計データの二次的利用」（調査票情報等の提供）に係る課題と今後の方向性 . . . P. 3
2. 「オンサイト利用」及び「プログラム送付型の集計・分析」
 - 2-1 オンサイト利用に関する仕組みの整備に当たっての論点 . . . P. 4
(第 14 回「統計データの二次的利用の促進に関する研究会」(平成 24 年 12 月 19 日)資料)
 - 2-2 前回の議論を踏まえた論点整理（オンサイト利用） . . . P. 8
(第 15 回「統計データの二次的利用の促進に関する研究会」(平成 25 年 3 月 26 日)資料)
 - 2-3 「オンサイト利用」及び「プログラム送付型の集計・分析」の実用化に
向けた課題整理と今後の進め方 . . . P. 10
(第 16 回「統計データの二次的利用の促進に関する研究会」(平成 25 年 6 月 25 日)資料)
3. 「統計データ・アーカイブ」の在り方についての検討状況 . . . P. 13
 - 3-別紙 1 「統計データ・アーカイブ」の今後の検討方向について . . . P. 14
(第 16 回「統計データの二次的利用の促進に関する研究会」(平成 25 年 6 月 25 日)資料)
 - 3-別紙 2 「公文書等の管理に関する法律施行令」（抜粋）及び「行政文書の管理に
関するガイドライン」（抜粋） . . . P. 19
 - 3-別添 公共データと「統計データ・アーカイブ」（イメージ図）(未定稿) . . . (別添)
- 参考 1 事業所母集団データベース（ビジネスレジスター）の概要 . . . (別添)
(第 39 回「統計委員会基本計画部会」(平成 25 年 6 月 27 日)総務省統計局説明資料（抜粋）)
- 参考 2 統計法令参照条文（調査票情報等の提供、事業所母集団データベース関係） . . . (別添)

「統計データの二次的利用」(調査票情報等の提供)に係る課題と今後の方向性

《検討に当たったての共通的な問題意識》
 提供する統計データに求められるセキュリティレベルごとに利用できる者・利用条件を設定する必要があるのではないか。

利用形態	法的根拠 〔法：統計法 省令：統計法施行規則〕	利用できる者・利用条件	求められるセキュリティレベル	課題と今後の方向性
調査票情報の提供 (公的機関以外の者が利用する場合)	法第33条第2号、 省令第9条、等	○公的機関が委託又は共同して調査研究を行う者 ○公的機関が公募の方法により補助する調査研究を行う者 ○行政機関等が政策の企画・立案、実施又は評価に有用であると認める統計の作成等を行う者	高	【課題】 利用者の研究室等で調査票情報の利用が可能。しかし、実地監査を行わない限り、実際にセキュリティが確保されているか確認できない。 また、利用開始前に利用場所、利用する環境、保管場所及び管理方法、作成予定の集計様式・分析出力様式を調査者実施機関が審査する必要があるが、申出・審査に係る利用者及び調査実施機関双方の負担が大きい。 ↓ 【今後の方向性】 ○調査対象者の秘密保護のためのセキュリティ確保に万全を期すとともに、利用の申出・審査に係る事務を効率化するため、オンライン利用やプログラム送付型集計・分析を推進
匿名データ	法第35条・36条、 省令第15条、等	○一般の者(学術研究目的、高等教育目的、国際比較目的)	中	【課題】 匿名データが提供されている統計調査の種類に限られている(平成25年6月末現在で6統計調査が提供中。また、25年内に国勢調査が追加予定。) また、提供中のものについても、新しい年次の追加要望がみられる。 ↓ 【今後の方向性】 ○匿名データの種類の追加 ○既に匿名データを作成している統計調査の年次の追加
委託による統計の作成等(いわゆる「オーダーメイド集計」)	法第34条、 省令第10条、等	○一般の者(学術研究目的、高等教育目的)	低	【課題】 いわゆる「オープンデータ」化の中で、行政機関が保有するデータに関し、とりわけ個人情報等の機微な情報を含まないものについては、営利目的を含めた民間分野での幅広い利用が求められている。 また、人手による作業が多くを占めていることから、オーダー(委託)を受けてから集計結果を提供するまでに時間を要するケースが多い。 ↓ 【今後の方向性】 ○利用制限の緩和 ○オンデマンド集計(インターネット上のシステムを利用した「オーダーメイド集計」のリアルタイム提供)に関する技術的な検証

利用に際しての申出・審査については、可能な限り事務の効率化、簡素化を図る必要がある。
 また、新たな利用形態の追加や既存の利用形態の拡充、利用条件の変更等を行う場合、ニーズや行政コストを勘案して料金設定を検討する必要がある。

オンサイト利用に関する仕組みの整備に当たっての論点

平成24年12月19日

総務省政策統括官室(統計基準担当)

1. 現状

- 日本においては、高度の公益性が認められる統計的な目的であれば、申請時に利用場所、利用する環境、保管場所及び管理方法等を申請書に記載した上で、利用者の研究室等で調査票情報を利用することが可能となっている。
- 一方で、利用に当たっては、利用目的の公益性や調査対象の秘密保護等の確保の観点から、調査実施機関による審査が行われているが、原則として、利用申請の段階で作成予定の集計様式や分析出力様式をすべて示すことを求めており、利用できる情報(調査事項)についても、集計様式や分析出力様式から見て必要最小限の範囲とされている。この点については、利用者及び調査実施機関双方の負担が大きいとの指摘がある。
- 統計的研究における調査票情報の活用の一層の推進のためには、調査対象の秘密保護のためのセキュリティ確保とともに利用の申請及び審査に係る事務の効率化が求められる。
- 現在、(独)統計センターでは、連携協力協定を締結している学術研究機関等(サテライト機関)のうち、国立大学法人一橋大学及び大学共同利用機関法人情報・システム研究機構の2か所をオンサイト利用施設と位置付け、総務省統計局と一橋大学の共同研究(「オンサイト利用施設の試行運用」として、2名の研究者が、一橋大学の施設を利用して、オンサイト利用施設の運用面での課題検討も含めた研究を実施中(研究途中の段階)となっている。
- 一方で、現行の「統計法第33条の運用に関するガイドライン」においてはオンサイト利用に関する記載がある(※)ものの、内容が不十分となっていることから、制度面の関係は必ずしも十分に整理されていない。
- なお、諸外国においては、調査票情報レベルのデータを利用させる場合、情報漏えい対策が十分に講じられたオンサイト利用施設を調査実施機関が指定し、そこに限定して利用させることが一般的となっている。(別添 参考①参照)

(※) 場所及び機器の指定に当たり、以下の点に留意する旨の記載等がある。

- ・ 情報管理の厳格な施設及び機器(物理的及びICT面のセキュリティが担保された作業環境及びデータ保管環境を整えた設備を有した施設・機器)であること
- ・ データ保護管理のルールを定め、施設及び機器の管理責任者、施設利用者の管理を行う利用管理者を配置した体制により管理されていること
- ・ 管理責任者、利用管理者により、利用者の監視措置、入退室、使用機器の調査票情報使用時における外部ネットワークとの遮断、利用者による不正な持込み及び持ち出しの防止などの措置が取られていること

2. 今後の整備に当たっての主な論点

(1) 意義・利点

オンサイト利用の仕組みを整備することにより、どのような意義・利点があるか

- (例)・ 利用者の研究室等では、使用環境の確認等の監査を行わない限り、実際に申請書どおりの利用環境が確保できているかどうか確認ができなかったが、オンサイト利用施設内ではセキュリティが確保されていることが担保されるのではないか。
- ・ オンサイト利用方式に移行するにつれ、研究室等での利用が減少し、利用環境のセキュリティ確保の実効性が向上するのではないか。
 - ・ 事前チェック方式(※1)から事後チェック方式(※2)に移行することにより、利用者側は施設内で試行錯誤的な研究が可能となるのではないか。また、事後チェックを効率的に行うことにより、提供者側の審査負担の軽減につながる可能性もあるのではないか。
 - ・ 現在は、各調査実施機関がそれぞれ所管する統計調査の調査票情報の提供を行っているが、オンサイト利用施設において一元的に提供を可能とすることで、利用者の利便性が向上するのではないか。
 - ・ 将来的には、調査票情報の提供以外の利用方法に活用できないか。

(※1) 使用する調査事項・作成する統計等を事前に確定して承認を得る方式

(※2) 施設内では試行錯誤的な研究ができ、施設外に成果物を持ち出す際に承認を得る方式

なお、この検討に当たっては、統計法第33条に基づく現在の調査票情報の利用の実態をできる限り把握し、比較検討を行う必要があるのではないか。

(2) 運用主体・運用形態等

- ・ オンサイト利用施設は誰が整備・運用するのか(行政機関、独立行政法人、委託を受けた公的研究機関等)
- ・ 各府省が委託する場合、具体的な委託内容・要件について、ガイドライン等により明確化・統一化を図る必要があるのではないか。また、施設の指定は一括して行う方が効率的ではないか。
- ・ 責任主体等を明確化するため、委託契約等のひな形を作成する必要があるのではないか。

(3) オンサイト利用施設の指定や運用に当たって整理すべき点

- ・ 施設管理者に求められる要件(研究者の支援、事後チェック能力等)について、整理する必要があるのではないか(ヒト)
- ・ 現行のガイドラインよりも具体的な施設・機器の情報漏えい対策等の必要要件を整理する必要があるのではないか(モノ)
- ・ オンサイト施設の運用コストを誰が負担すべきか(施設使用料を徴収することについてどう考えるか)(カネ)

なお、行政資源に限りのある中で、利用者の利便性も確保しつつ、オンサイト利用施設を運用していくためには、公的研究機関等の理解と協力を求めていく必要があるのではないか。

諸外国における統計データの二次的利用制度の状況(概要)

(参考①)

提供方法		提供データの種類		
	調査票情報レベルのデータ	匿名データ	パブリックユース ファイル	
直接利用型	オンサイト型	アメリカセンサス局、 アメリカ保健統計センター、 イギリス、韓国(政府から委託を受けた 者のみ利用可)	カナダ、ドイツ、オランダ、 オーストラリア、 ニュージーランド(加工度低)	—
	直接利用型	日本	イギリス、ドイツ、オーストラリア、 ニュージーランド(加工度高)、 日本	アメリカセンサス局、 アメリカ保健統計センター、 カナダ、ドイツ、韓国
プログラム 送付集計型	参照可能型	韓国(政府から委託を受けた者のみ利用 可)	オランダ、スウェーデン、 ニュージーランド(加工度中)	—
	参照不可能型	アメリカ保健統計センター、 ドイツ(国内研究者向け)	カナダ、ドイツ(国外研究者向け)、 オーストラリア	—
オナーメー ド型	従来型 (後日提供型)	アメリカセンサス局、カナダ、 イギリス、オーストラリア、 ニュージーランド、韓国、 日本	—	—
	リアルタイム 提供型	オランダ、オーストラリア (集計元データに、データキューブを利 用しているものを含む)	—	—

(注)1 韓国では、調査票情報は、政府から委託を受けた者のみ利用可能となっている。

2 ニュージーランドでは、匿名データは、利用方法に応じて加工度(秘匿度)が異なっている。

3 ドイツでは、プログラム送付集計型による提供は、国内研究者向け(調査票情報レベルのデータ)と
国外研究者向け(匿名データ)とで集計元情報が異なる。

4 オランダ及びオーストラリアでは、オナーメード型(リアルタイム提供型)について、一部で集計元データにデータキューブ
と呼ばれる調査票情報を用いて作成された高次元クロス集計表を用いている。

5 これら以外に、統計教育目的にレプリカデータ等を提供している例がある。(次頁参照)

統計教育サービスの状況

調査対象機関名	サービス名	サービス内容
ドイツ連邦統計局	Campus Files	統計教育のために作成された匿名データの種類。無料で提供されている。このファイルは特別に教育目的に作成されているもので強い秘匿が行われており、詳細な分析に向いていないとしている。より秘匿性の低い匿名データを利用する前の訓練用という位置付けとなっている。
ニュージーランド統計局	SURFs (Synthetic Unit Record File) の提供	統計教育に作成されたレプリカデータ。公表されている集計結果から超多次元表を作成し、繰り返し比例補正の手法を用いて作成したものである。無料であり、ウェブサイトを通じて誰でもダウンロードすることができる。
韓国統計庁	統計教育用マイクロデータ	特に大学生及び大学院生のために作成されたマイクロデータ。無料でWeb上に構築されたMDSS (Micro Data Service System) 経由でダウンロードできる。

前回の議論を踏まえた論点整理（オンサイト利用）

平成25年3月26日

総務省政策統括官室（統計基準担当）

前回（平成24年12月19日）の研究会での主な意見と今後の方向性について、以下のとおり整理を行った。

（オンサイト利用の意義・利点）

- オンサイト利用については、セキュリティの確保を前提に、調査実施者と利用者の双方にメリットがあるような使い方を目指すべき。（廣松座長）
- 調査票情報の利用手続の簡素化と利便性の向上に資するため、オンサイト利用施設内では探索型の利用ができるようにし、成果物について個人情報漏れることがないかを厳格にチェックする方向で「統計法第33条の運用に関するガイドライン」（平成20年12月24日総務省政策統括官（統計基準担当）決定。以下「ガイドライン」という。）を改正することを検討。
また、その際、複数の統計調査に係る調査票情報のリンケージデータの活用や分析用プログラムの共有等の研究手法の高度化・効率化の方法についても、併せて検討。

（施設の設置に当たって整理すべき点）

- オンサイト利用施設が広がっていく際、地理的な条件や利用者の条件について、公平性の観点から踏み込んだ議論が必要ではないか。（玄田委員）
- 何らかの形で大学等の公的機関との協力は不可欠。特に地方にある大学等の公的機関との協力や契約が必要と思われる。（廣松座長）
- 公平性という観点では、手を挙げた機関は全部採択するというのも一つの方法。基本的には、利用者が多くいると思われる大学等が希望し、立候補した場合には拒否せず認めるという方法がいいのではないか。また、昔のいわゆる大型計算機センターを活用するのも一つの方法（この場合、利用料金を取ることを認めることが必要。）。現在でも各大学で施設・設備を有しており、一部屋でも利用できればオンサイト利用施設として運用できるため、基準をきちんと整備すれば公平性の観点からも割と機能するのではないか。（安田委員）
- 施設利用料の徴収については、認める方向で考えて良いのではないか。また、既にデータを管理・利用するための施設があれば、それを活用することが効率的ではないか。（廣松座長）
- 利用機会の公平性を確保するため、オンサイト利用施設の地方拠点を設置することが必要であり、地方の大学等との連携が必要。また、ある時点で急に現行の運用を凍結させることは適切ではないが、情報漏えい対策の観点からは、地方拠点の整備状況に併せて徐々にオンサイト利用施設での利用にシフトさせるといった措置が必要。
- 施設利用料については、アメリカ・カナダなどの考え方を参考としつつ、日本における基本的な考え方を議論すべき。

(現行の運用における利用実態の把握)

- 議論の前提として、利用実態の把握は大変重要。(廣松座長)
- 今後、統計法第 33 条に基づく調査票情報の利用実態について、各府省と情報共有を進める。

(運用主体・運用形態に関する共通基準の設定)

- 今後、施設を立ち上げることを考える場合、運用主体や運用形態について、各府省共通の基準を設けていただけるとありがたい。利用者側・提供者側の双方にとって有益であるという印象。(樫委員)
- 運用主体や運用形態について、アメリカ・カナダなどの考え方を参考としつつ、日本における基本的な考え方を議論すべき。研究会での議論の結果を踏まえ、ガイドラインの改正や各府省共同での方針の申合せなどの方法で対応する方向で検討。

(承認の統一基準)

- 現行制度では、各統計調査の実施者ごとに承認を受けることとされているが、ガイドライン等により基準の統一化を図った方が良いのではないか。(安田委員)
- 基本的審査基準については現行のガイドラインに記載があるものの、さらなる具体化について要検討。また、アメリカ・カナダなどの現状を参考としつつ、成果物の事後チェックの方法についても議論を深める必要(様々な統計調査ごとの特性に精通した者によるチェックが必要であり、現実的には、現行の運用と同様、調査実施者による審査が必要ではないか。)

(オンサイト利用施設管理者の位置付け)

- オンサイト利用施設において問題が発生した場合の施設側の管理責任について、ペナルティを含め、踏み込んだ形で議論する必要があるのではないか。また、現行制度では利用者でない限りはオンサイト利用施設の管理者であっても調査票情報を利用できないため、この点についても整理が必要。(安田委員)
- 一橋大学での試行運用においては施設管理者を 33 条の申請者(個々の利用の共同研究者)として取り扱っているが、施設管理者の位置付けを明確にした上で、施設管理者側の管理責任等について検討。

(利用目的が拡大された場合の対応)

- 将来的に二次的利用の利用目的が拡大された場合、大学等の学術研究機関に設置された場合でも学術研究以外の目的に開かれるということが必要ではないか。(縣委員)
- 大学も社会的貢献が求められており、利用目的の拡大があった場合に利用の門戸を広げることは大学側の方針にも沿ったものと言えるのではないか。(廣松座長)
- 今後利用目的が拡大された場合、要検討。

(その他：プログラム送付集計型)

- 以前プログラム送付集計型の研究をしたことがあるが、利用者の技能不足が原因で送付されたプログラムが動かないことが多かった。プログラム送付集計型を採用する場合、事前研修も含めた利用者の訓練を行うことが必須。(安田委員)
- 発展的・応用的な課題として、今後も引き続き検討。

「オンサイト利用」及び「プログラム送付型の集計・分析」の 実用化に向けた課題整理と今後の進め方

平成25年6月25日

総務省政策統括官（統計基準担当）

1 背景

前回（平成25年3月26日）の研究会において、プログラム送付集計は発展的・応用的な課題としていたが、オンサイト利用施設への物理的なアクセスが難しい地方在住の利用者の利便性確保の手段として、プログラム送付型の集計・分析は有力な解決策の一つであると考えられる。

そこで、当研究会において、オンサイト利用に加え、プログラム送付型の集計・分析の実用化に向けた課題整理も行うこととしたい。

2 課題整理

（1）オンサイト利用

前回までの議論において、情報漏えい対策の観点から、現行の調査票情報の利用形態からオンサイト利用施設での利用に徐々にシフトし、オンサイト利用施設内では探索型の利用ができるようにし、成果物を持ち出す際に個人情報漏れることがないかを厳格にチェックするという方向性について、おおむね共通認識が得られたものと考えられる。この方向性に進めていく上での具体的な課題について、再度整理した。

○ 運用形態（「オンサイト利用に関する提案」平成24年12月19日 資料1-2から引用）

→ ①直轄方式：政府統計機関またはその委託を受けた独立行政法人（以下「政府統計機関等」という。）が、その建物内に設けたオンサイト施設の運用を行う方式。オンサイト施設設置機関は政府統計機関等である。

②委託方式：政府統計機関等からオンサイト利用の運用業務の委託を受けた大学法人または大学共同利用機関法人（以下、「運用受託機関」）が、その建物内に設けたオンサイト施設の運用を行う方式。オンサイト施設設置機関は運用受託機関である。

※ 業務委託は法人間の契約

・運用受託機関内でオンサイト施設の運用を実際に担当する部局・部署を明確にする。

※ 契約に当たっては、

・施設が技術的基準を満たさなくなったり、運用面で不都合が生じたりした場合は改善がみられるまで業務停止。

・組織が調査票情報等の不適切使用などを看過したりした場合は認定取消を条件とする。

○ 施設利用料

→ アメリカセンサス局では、利用料を徴収しているものの運営経費を賄うには至らず、保健統計センターでも運営費全体の5～10%程度しか賄えず、ハードウェアの購入に充てられる程度であり、興味本位での利用をけん制する機能としての意味合いが大きい。

カナダでは、大学がオンサイト利用施設を運営しており、利用料を取るかどうかは大学に委ねられている。

諸外国の例を参考としつつ、これまでの経緯等を踏まえ、利用料設定の考え方について、引き続き検討を進める。

○ 施設内で作成した分析用プログラムの共有

→ 公的統計のデータは、国民の共有財産として、幅広く用いられるべきものであることから、オンサイト利用施設内で作成した分析用プログラムは共有する方向で検討。

○ 利用機会の公平性の確保

→ オンサイト利用の推進と併せて、プログラム送付型の集計・分析サービスを提供することにより、地方在住の利用者のニーズを満たすこととしてはどうか。

(2) プログラム送付型の集計・分析

諸外国における実用例を参考に、我が国においても統計法第33条（調査票情報の提供）の一形態とし、プログラム送付型の集計・分析の導入の検討が必要と考えられるが、運用面では以下の課題がある。

○ テストデータの提供

→ 利用者側は、実際のデータに近い分布を保持し、データレイアウトも実際のデータと同一なテストデータの利用を求めるものと考えられ、また、プログラム送付型集計・分析サービスの提供者側としても、きちんと実行されるプログラムを受領することが望ましい。

→ 多くの統計調査で早期にプログラム送付型の集計・分析を利用可能とするため、架空のダミーデータを新たに作成するのではなく、調査票情報からリサンプリング等の必要な措置（ただし、項目の削除や符号の統合は行わない。）を施したテストデータを作成し、必要なセキュリティ対策を措置された環境下において、それを利用する方法が考えられる。

→ ドイツ連邦雇用庁では、オンサイト利用が原則であり、滞在期間が終了しても集計・分析が終わらなかった場合にプログラム送付型の集計・分析を認める形となっている。すなわち、オンサイト利用施設で研究者が実際にミクロデータを利用することが、プログラム作成の訓練となっている。

なお、（独）統計センターが試行提供を行っている擬似ミクロデータについては、研究開発に相当程度の期間が必要であり、また、既に作成・提供済みの全国消費実態

調査の擬似マイクロデータについても、擬似マイクロデータとして提供されている項目が限定されており、元のデータと同一のデータレイアウトになっておらず、プログラム送付型の集計・分析用のテストデータに用いることは困難と考えられる。

○ 集計・分析結果の提供までに要する期間

→ 調査実施者側は、結果の秘匿性が確保されているかをチェックするため、結果提供までに相応の時間を要する。一方、利用者としては、回帰分析等の統計分析において変数の入れ替えを行う場合など、様々な試行錯誤を行いたいため、結果提供までの期間はできる限り短い方が望ましい。

→ 多数の集計表が作成された場合、秘匿性のチェックに時間がかかると考えられ、できる限り短期間で結果提供を可能とする方法（例：利用者が作成するプログラムに予め必要な秘匿性チェックのスキームを組み込むなど）を検討する必要がある。また、統計分析の結果のチェックには比較的時間を要しないと考えられるが、元のデータを復元できないようにする必要がある。

○ 利用可能なデータ

→ プログラム送付型の集計・分析の対象とするデータについては、調査対象の秘密保護等の観点やプログラムの作成・実行が容易かどうか等の観点から検討する必要がある。

(3) その他（利用者教育）

統計データの二次的利用の促進を図るため、大学等の高等教育機関におけるマイクロデータの取扱いに関する教育を推進する必要がある。その一環として、(独)統計センターで試行提供を行っている擬似マイクロデータについて統計調査の種類を追加する等、教育用マイクロデータの一層の開発・提供が必要ではないか。

3 今後の進め方

政策統括官（統計基準担当）は、本研究会において整理したオンサイト利用に必要な要件の具体化を進める。

また、(独)統計センターにおいて、プログラム送付型集計・分析の実用化に向けた技術的研究を進める。

「統計データ・アーカイブ」の在り方についての検討状況

平成 25 年 7 月 9 日

総務省政策統括官（統計基準担当）

1 検討経緯

- 「統計データの二次的利用の促進に関する研究会」（以下「研究会」という。）において、「統計データ・アーカイブ」について、海外事例の検討及び我が国における整備・運用に当たっての議論を実施。
- その結果、現行基本計画の記述自体に多義性があり、我が国における整備・運用方法を導き出す上で整理しておくべき様々な課題があることが判明。

2 検討状況

- 「統計データ・アーカイブ」に期待される機能として、（１）統計データの収集・整理・保管、（２）統計機関相互のデータ共有・連携、（３）ユーザーへの提供の三つに大別し、それぞれについて「視点」及び「検討の焦点」を整理。（別紙 1）
- 研究会において、現在の分散型統計制度の前提条件や人員・予算等の環境を踏まえると、当面、「統計データ・アーカイブ」固有の課題として「統計データの適切な収集、整理、保管の確保」に注力するとの方針について了承。
- その他の「視点」については、例えば「ユーザーへの提供」に関してはいわゆる「二次的利用」に係る取組を進め、その実績を踏まえて整理するなど、「統計データ・アーカイブ」固有の課題としては取り扱わないものの、別途の枠組みの中で検討を進め、将来的には「統計データ・アーカイブ」機能に統合。
- 現在、政府全体としてオープンデータ政策の取組を推進しており、当該政策の動向も注視する必要。

3 今後の方向性

以下の方向性について、次期基本計画案に盛り込むことが適当と考える。

- 統計データの二次的利用促進のため、統計データ・アーカイブの整備を進める。特に、統計データとともに収集、整理、保管すべきメタデータ^(※1)について整備を進める。

(※1) データそのものではなく、そのデータに付随するデータ自身についての関連情報。データの作成日時や作成者、データ形式、タイトル、注釈などが考えられ、データを効率的に管理したり検索したりするために重要な情報である。また、統計データの活用に関しては、データの内容・特性の理解を助けるため、実査や集計等の統計作成の各段階における作業がどのように行われたかについての情報（調査時期、調査方法、調査対象、抽出方法、推計方法等に関する情報）を整備することが重要。

なお、調査票原票の保管^(※2)の必要性については、二次的利用の観点や歴史的公文書としての価値の観点から検討を行い、取扱いを定める必要があるのではないかと。

(※2) 現行では、概ね数年以内に設定された保存期間終了後に廃棄されている。（公文書の保存期間や保存期間満了時の措置の設定基準については、別紙 2 参照。）

「統計データ・アーカイブ」の今後の検討方向について

平成25年3月26日

平成25年6月25日(修正版)

総務省政策統括官(統計基準担当)

前回の研究会において委員から頂いたご意見等を踏まえ、「統計データ・アーカイブ」について改めて検討・整理を行った。

【第一段階】

「統計データ・アーカイブ」に期待されている複数の機能とそれぞれの機能に関する視点を整理し、当該視点ごとに、現状や環境、ニーズなどを総合勘案した上で焦点を設定する。

(概要)

これまでの検討により、「統計データ・アーカイブ」に期待されている機能は統計データの利用支援を目的とする収集・整理・保管、統計機関相互のデータ共有・連携、ユーザーへの提供に大別できるのではないかと認識された。このそれぞれについて基本計画や当研究会における調査や検討の結果、様々な視点があることが認識された。これを踏まえ、議論を拡散させずに検討を進めるために、視点ごとに焦点を整理する。(イメージ 別添参照)

→ 抜け落ちている視点はないか。視点ごとの焦点の設定は適切か。

【第二段階】

第一段階で設定した焦点について、平成25年度中に行う基本計画の変更の際に、当該計画に「統計データ・アーカイブ」に係る具体的な取組を盛り込むことを目標として、当面検討を深めるものを絞り込む。

→ 検討対象の選別(絞り込み)を実施。

(別添)「統計データ・アーカイブ」に期待される機能に係る視点の整理と今後の検討の焦点及び方向性

(1) 統計データの収集、整理、保管

機能	視点	検討の焦点	方向性
a 収集、整理、保管	a-1 一元化された専門機関の設置	○ 機関設置の意義・目的の整理と現行制度との関係の整理ではないか。	○ 現在の分散型統計システムを前提とすれば、機関の活動の対象は(現用のデータではなく)史料価値のある統計データであるとして構想することが自然。典型的なアーカイブ機関の構想として論ずることになるのではないか。 ○ 機関設置は短期間で成案に至る環境ではないのではないか。理想的将来像の一つとしての位置付けになるか。 ○ 統計データを適切に収集、整理、保管すること自体の重要性は自明。(a-2)の視点へ。
	a-2 適切な収集、整理、保管の確保	○ 現在の収集、整理、保管に関する取組の問題点と対処方策ではないか(収集、整理、保管)。 ○ <u>対象データの範囲(調査票情報そのものか、統計データのいわゆる「二次的利用」のために作成した匿名データ等か)</u>	○ アーカイブ機関による一括管理を将来像の一つとするのであれば、各行政機関の取組についてどのように統一性の確保をすべきか検討するのではないか。 ※ なお、「調査票情報等の管理及び情報漏えい等の対策に関するガイドライン」(平成23年3月28日付け総務省政策統括官(統計基準担当)決定)で、電磁的に記録した調査票情報(統計データ)や関連ドキュメント(メタデータ)については永年保存として管理する方針が示されているように、アーカイブ以外のアプローチからもこの問題への取組がなされていることに留意。

→ 検討対象の選別(焦点の絞り込み)

○ 「a-1」については、専門機関の設置は将来の検討課題とすることが適当ではないか。

○ 「a-2」については、統計データ(及びメタデータ)の収集、整理、保管に関する運用の統一性については一定程度確保されているものと考えられるが、情報技術の進展状況に応じて、保存媒体の見直し等を行うことが必要ではないか。また、統計データのより一層の有効活用に資するため、特にメタデータの整備について検討を進める必要があるのではないか。

なお、調査票原票自体の保管については、物理的な場所の確保等、多大のコストを要することが見込まれ、原票を保管対象とすることは現実的ではないと考えられる。

(2) 統計機関相互のデータ共有・連携

機能	視点	検討の焦点	方向性
b 共有・連携	b-1 共有による活用、特に統計調査の重複是正	<ul style="list-style-type: none"> ○ 当初目的以外の行政機関による統計データ活用の環境整備 ○ 本格運用を開始した「事業所母集団データベース」の活用をいかに進めるかの問題が大きいのではないか 	<ul style="list-style-type: none"> ○ 統計データ・アーカイブ固有の問題ではないのではないか。メタデータ整備、データ形式の統一等の課題であれば、a-2の視点に係る検討の一環として扱ってよいのではないか。他方、事業所母集団データベースの活用については、当該データベースの課題として整理してはどうか。
	b-2 連携による国民への統計データの提供のワンストップ化	<ul style="list-style-type: none"> ○ 現在、e-Stat によりかなり実現しているが、その上に必要なニーズは何か 	<ul style="list-style-type: none"> ○ 統計データ・アーカイブの固有の問題ではなく、統計表の提供については e-Stat の課題、その他は統計データのいわゆる「二次的利用」の促進の取組として整理してはどうか。 (ちなみにオーダーメイド集計及び匿名データの提供については、各府省において受付を実施しているものがあるほか、事務が全部委託された統計調査に関するものについては、(独)統計センターに国民への提供の窓口の一元化が図られている。)

→検討対象の選別 (焦点の絞り込み)

- 「b-1」については、メタデータ整備、データ形式の統一等の課題については「a-2」の視点に係る検討の一環として扱い、事業所母集団データベースの活用については当該データベースの課題として整理してはどうか。
- 「b-2」については、統計表の提供については e-Stat の課題、その他は統計データのいわゆる「二次的利用」の促進の取組として整理してはどうか。

(3) ユーザーへの提供

機能	視点	検討の焦点	方向性
c 提供	c-1 セキュリティの十分な場所での提供の確保	<ul style="list-style-type: none"> ○ 現在検討しているオンサイト利用等についてどう考えるか。 	<ul style="list-style-type: none"> ○ オンサイト利用の仕組みの構築はそれ自体を一つの重要課題として考えてよいのではないか。 ○ <u>オンサイト利用に加え、実際のデータには直接触れないプログラム送付型の集計・分析についても検討すべきではないか。</u> ○ 統計データ・アーカイブという論点から見る限り、アーカイブによる将来の「統計データの提供」は、オンサイト利用等の実績を踏まえて考え方を整理することとすべきではないか。

	c-2 データの格付け	<p>○ 提供する統計データに求められるセキュリティレベルごとに、提供する際の目的や資格に関する制約を柔軟にする必要があるのではないか。</p>	<p>○ 統計データ・アーカイブの固有の問題ではなく、統計データのいわゆる「二次的利用」の促進に係るものとして扱ってよいのではないか。</p> <p>○ 統計データ・アーカイブという論点から見ると、アーカイブによる将来の「統計データの整理（及び提供）」は、二次的利用における取組の実績を踏まえて考え方を整理することとすべきではないか。</p>
--	-------------	--	---

→検討対象の選別（焦点の絞り込み）

- 「c-1」については、オンサイト利用の仕組みの構築の検討を進めるとともに、プログラム送付型の集計・分析についても検討を行うこととし、その実績を踏まえて整理することとしてはどうか。
- 「c-2」については、まずは統計データのいわゆる「二次的利用」に係る取組を進め、その実績を踏まえて整理することとしてはどうか。

【参考】「公的統計の整備に関する基本的な計画」（平成 21 年 3 月 13 日閣議決定）
（統計データ・アーカイブ関係部分の抜粋）

第 1 公的統計の整備に関する基本的な方針

3 施策展開に当たっての基本的な視点

（3）統計データの有効活用の推進

（略）

匿名データ等有効なデータを蓄積し、学術研究等の目的での匿名データ等の利用の便を図るための基盤として、統計データ・アーカイブ^{（注 1）}を整備することも必要である。

（注 1）統計調査の調査票情報から作成された匿名データ等を収集、整理、保管し（統計調査の調査票情報を含める場合もある。）、学術研究等の目的で匿名データ等を提供する機関をいう。

第 3 公的統計の整備を推進するために必要な事項

4 統計データの有効活用の推進

（2）統計データ・アーカイブの整備

イ 取組の方向性

限られた統計リソースの効率的、効果的な活用を図る観点から、統計データ・アーカイブは、基本的には一つの機関に集約することとし、この機関にどのような機能を持たせるか、調査票情報まで蓄積すべきか等の詳細については引き続き検討する。その際、政令指定法人、学会等の協力を得て、統計データ・アーカイブの検討を行う。

別表：今後 5 年間に講ずべき具体的施策

○ 統計データ・アーカイブの整備に向け、以下の取組を実施する。

- ・ 各府省、統計センター、学会等の協力を得て、検討会議を設置し、統計データ・アーカイブの整備・運用方法、保有すべき機能、対象データの範囲や保存方法を検討し、結論を得る。
- ・ 調査票情報の提供、オーダーメイド集計並びに匿名データの作成及び提供の将来の在り方についても併せて検討し、結論を得る。

担当府省：総務省

実施時期：平成 25 年度までに結論を得る。

公文書等の管理に関する法律施行令（抜粋）

（公布 平成 22 年 12 月 22 日 政令第 250 号、最終改正 平成 23 年 12 月 26 日政令第 421 号）

（行政文書ファイル等の分類、名称及び保存期間）

第 8 条 第 2 項

法第 5 条第 1 項の保存期間は、次の各号に掲げる行政文書の区分に応じ、それぞれ当該各号に定める期間とする。

- 一 別表の上欄に掲げる行政文書（次号に掲げるものを除く。） 同表の下欄に掲げる期間
- 二 他の法律又はこれに基づく命令による保存期間の定めがある行政文書 当該法律又はこれに基づく命令で定める期間
- 三 前 2 号に掲げる行政文書以外のもの 別表の規定を参酌し、行政機関の事務及び事業の性質、内容等に応じて行政機関の長が定める期間

行政文書の管理に関するガイドライン（抜粋）

（平成 23 年 4 月 1 日 内閣総理大臣決定、平成 24 年 6 月 29 日一部改正）

別表第 2 保存期間満了時の措置の設定基準

- 1 （略）以下の【Ⅰ】～【Ⅳ】のいずれかに該当する文書は、「歴史資料として重要な公文書その他の文書」に当たり、保存期間満了後には国立公文書館等に移管するものとする。

- 【Ⅰ】国の機関及び独立行政法人等の組織及び機能並びに政策の検討過程、決定、実施及び実績に関する重要な情報が記録された文書
- 【Ⅱ】国民の権利及び義務に関する重要な情報が記録された文書
- 【Ⅲ】国民を取り巻く社会環境、自然環境等に関する重要な情報が記録された文書
- 【Ⅳ】国の歴史、文化、学術、事件等に関する重要な情報が記録された文書

2 具体的な移管・廃棄の判断指針

1 の基本的考え方に基づいて、個別の行政文書ファイル等の保存期間満了時の措置（移管・廃棄）の判断については、以下の(1)～(4)に沿って行う。

- (2) 以下の左欄の業務に係る歴史公文書等の具体例は、右欄のとおりであることから、これらの歴史公文書等を含む行政文書ファイル等に移管することとする。

業務	歴史公文書等の具体例
各行政機関において実施・運用している制度（例：政策評価、情報公開、予算・決算、補助金等、機構・定員、人事管理、統計等）について、制度を所管する行政機関による当該制度の運用状況の把握等の業務	<ul style="list-style-type: none"> ・基本計画 ・年間実績報告書等 ・施行状況調査・実態状況調査 ・意見・勧告 ・その他これらに準ずるもの
(略)	
統計調査	<ul style="list-style-type: none"> ・基幹統計調査の企画に関する文書及び調査報告書 ・一般統計調査の調査報告書