

学術研究の発展や EBPM の促進を図るための公的統計の調査票情報の提供円滑化について（論点）

令和 5 年 3 月 6 日  
事 務 局

## 1. 基本的な考え方

○学術研究の発展、EBPM の促進を図るためには、公的統計の統計調査に対する国民の信頼や協力を確保しつつ、デジタル技術を最大限利用した、公的統計の調査票情報の利用環境の整備が必要不可欠。このため、以下のとおり、提供する調査票情報の範囲の限定のみではなく、利用者・機関の限定等の手法も利用して、個人や法人の権利・利益の保護を適切に図り、必要なリソースを確保した上で、今後数年以内に欧米に比して遜色がない利用環境を実現する必要があるのではないかと。

## 2. 磁気媒体の提供（申請及び審査手続）

### 【課題】

○ガイドライン（「調査票情報の提供に関するガイドライン」）において利用する調査事項（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（以下「必要最小限要件」という。）が求められており、省庁毎の審査担当者が目的の公益性および目的との関係で必要最小限か否かを審査することに膨大な時間を要するとの指摘がある。

○省庁ごと申請書の様式や記載方式が変わり、また、審査において説明が求められる内容が担当者によって異なる場合があるとの指摘がある。

### 【対応の方向性】

○調査票情報の提供に要する期間の大幅な短縮及び利用者にとっての予測可能性の確保、利用者の負担を合理的に最小限度とするために以下の措置を検討する必要があるのではないかと。

- ・提供に係る審査の標準化・効率化、進行管理や提供に時間がかかっている場合の相談対応の充実、リソースの確保を図る必要があるのではないかと。
- ・提供に係る審査の際のチェックリストの公表や審査の際に求める資料についての Q & A の整備・充実を図る必要があるのではないかと。
- ・上記の実務的な対応の他、研究開始前に「必要最小限」な変数を特定することは極めて困難である、研究開始後に探索的研究の過程で必要な変数を特定可能であるとの指摘を踏まえると、「必要最小限要件」の在り方について、運用を含め、抜本的な見直しを行い、遅くとも、申請後数週間以内に適切な調査票情報の提供が行われる制度・運用を確保する必要があるのではないかと。

いか。

- ・このためには、個人識別リスクが相当程度大きい変数を特定し（ブラックリスト作成）、当該変数に審査を重点化することも検討する必要があるのではないか。
  - ・個別申請の審査（公益性の審査、ブラックリスト変数など）におけるピアレビューの採用による専門性の確保、審査体制に応じた標準処理期間の設定、省庁横断的な一元審査体制の検討（※）などを含め大幅な効率化策を検討する必要があるのではないか。
- ※省庁毎、課室毎の現在の実務を踏まえ、「悪化」しない工夫が必要。

### 3. リモートアクセスの実現及びオンサイト施設利用の円滑化

#### (1) リモートアクセスの実現

##### 【課題】

○オンサイト施設の施設数、利用可能時間、同時利用可能人数が物理的に限られており、災害時や感染症流行時などには利用が困難である。また、オンサイト施設において集計等を行った結果を持ち出す際には、提供機関が個人識別リスクや不適切な利用を審査するため、他の研究者との共同研究（例えば、オンサイト施設を利用しない研究者がいる場合）が容易ではないとの指摘がある。

##### 【対応の方向性】

○欧米における統計データの利用環境も参考にしつつ、今後数年以内に基本的には全統計について研究室や行政機関などからのリモートアクセスを実現することとし、そのために必要なリソースを確保して準備を進める必要があるのではないか。その際、リモートアクセスには「必要最小限要件」を課さないこととし、例えば、一定の要件を満たした機関・研究者に提供するといった手法について検討し、円滑な共同研究の確保等を図る必要があるのではないか。

#### (2) オンサイト施設での利用円滑化

##### 【課題】

- オンサイト施設で利用可能な調査票情報は、その整備が追いついていないため、対象年次、対象統計の範囲の両面で媒体データの提供による場合よりも限定されている。
- オンサイト施設の利用要件として、過去に調査票情報の提供を受けたこと等過去の取扱実績が求められており、海外から帰国したばかりの研究者が調査票情報を利用することが困難な場合があるとの指摘がある。同様に、競争的資金を受けて行う調査研究において研究者として位置付けられているような場合を除き、学生が研究目的で調査票情報を利用することができない。

##### 【対応の方向性】

○オンサイト施設で利用可能な調査票情報の充実が必要であることから、そのためのリソースを

確保し、年限を区切り計画的に進めるための工程表を来年度早期に作成し、数年以内にオンライン施設での調査票情報の充実を完了する必要があるのではないか。

○上記と同様、一定の要件を満たした機関・研究者に提供するといった手法について検討する必要があるのではないか。

#### 4. その他の課題

##### (1) メタデータの整備

###### 【課題】

○利用可能な調査票情報の範囲や定義が分かりにくく、二次利用向けにどのような変数がどの統計に存在するのか十分に可視化されていないとの指摘がある（現状、ポータルサイトで一部のみ公開）。この結果、提供を受けた調査票情報が手元に来た際に想定と異なっていることや、わからないからとりあえず申請しておく変数が存在することがある。

###### 【対応の方向性】

○メタデータの整理（利用可能な調査票情報の年次、調査年ごとに利用可能なすべての変数のリスト、調査年ごとに利用可能な変数の定義）を、利用者の提供希望の有無に関わらず行っていくため、リソースを確保して計画的に実施するための工程表を来年度早期に作成する必要があるのではないか。

##### (2) 機械可読データの整備

###### 【課題】

○オープンデータを含め、機械可読性がない状態のデータ（PDF 等）が存在し、集計が困難との指摘がある（手で入力を行う必要）。

###### 【対応の方向性】

○機械可読性のないデータについては、早期に整備を行うため、リソースを確保して、年限を区切った工程表を来年度早期に作成する必要があるのではないか。

##### (3) 各種統計の連結解析に必要な紐付け手段の整備

###### 【課題】

○各種統計について、今後様々なデータベース（国民の所得情報等）との連結によって、学術研究やEBPMに大きな効果が生じることが期待されている一方、我が国では、連結に必要な識別子が存在しない。現状、氏名・生年月日・性別等を基に連結している事例も見られるが、限界も大きいとの指摘がある。

###### 【対応の方向性】

○連結するための方策について、諸外国の状況を踏まえ、検討が必要ではないか。