

## 第6回 点検検証部会 議事概要

1 日 時 令和元年6月13日（木）9:30～12:00

2 場 所 総務省第2庁舎6階特別会議室

3 出席者

### 【委員】

河井 啓希（部会長）、西郷 浩（部会長代理）、川崎 茂、嶋崎 尚子

### 【専門委員】

大西 浩史、西 美幸

### 【審議協力者】

内閣府大臣官房企画調整課課長補佐

総務省統計局総務課長

財務省大臣官房総合政策課経済政策分析官

文部科学省総合教育政策局調査企画課長

厚生労働省政策統括官付参事官付審査解析室長

農林水産省大臣官房統計部統計企画管理官

経済産業省大臣官房調査統計グループ統計企画室長

国土交通省総合政策局情報政策課長補佐

埼玉県総務部統計課長

東京都総務局統計部長

### 【事務局（総務省）】

横山大臣官房審議官、平野大臣官房審議官

統計委員会担当室：櫻川室長、永島次長、阿南次長、柴沼次長

政策統括官（統計基準担当）付：澤村統計審査官

4 議 題

（1）重点審議（学校基本調査（システム変更の柔軟性）について）

（2）重点審議（「プログラムミスによる統計数値の誤り発生の防止」の審議の進め方について）

（3）第1次再発防止策（案）について

（4）その他

5 概 要

（1）重点審議（学校基本調査（システム変更の柔軟性）について）

河井部会長から、本日の議題について説明。続いて、文部科学省から、資料1により学校基本調査システムの柔軟性の確保に向けた取組について説明が行われ、質疑が行

われた。

主な発言は以下のとおり。

- ・文部科学省には事前に質問項目をお渡ししているので、回答をお願いします。今日の審議でどこまで議論するか、委員各位と共通認識を持ちたい。開発がうまくいくように、どこに力点・チェックポイントを置くか、注意点を問題提起して、その後の対応は文部科学省に任せるということによいのではないか。
- ・新システムは6種類の調査全てを対象とすることを想定しているのか。  
→学校基本調査の移行を最優先にしつつ、他の調査も順次移行する。
- ・復元推計はどの段階で行うのか。そのための復元乗率はシステムのどこで算出されるのか。その場合、母集団情報が必要だが、どこから入手し、どのように保存するのか。  
→復元推計は集計段階のプログラムにおいて実装することを想定。母集団情報は、行政記録情報をマスターデータとしてデータベースに格納し、これを用いて復元乗率を算出する。
- ・データチェックは、過去データとの比較を行うのか。過去データはいつまでデータベース化するのか。過去にさかのぼって調査対象をたどれるよう、何からの統一的なキーを用いるのか。  
→過去データは保存期間が永年となっているものを対象にデータベースに格納予定。また、学校ごとに学校調査番号を付し、その番号を統一的なキーとする予定。
- ・回答された調査票情報がシステムによりチェックされ、エラーが発見された場合には、誰がどのようなルート、手続きにより確認、訂正を行うのか。確認、訂正の実行確認や記録はどのように行うのか。  
→システムで検出したエラーはリストとして出力され、経由機関で確認、訂正を行う。変更記録を残すことは提案を業者に募っていきたい。
- ・オンライン調査の場合、調査過程で経由機関は調査票内容を見ることがないのでは。だとすると、エラーチェック、修正に適切に関与できるのか。  
→オンライン調査システムで調査票情報の確認ができる。文部科学省のシステムでも途中段階の集計結果は確認可能。
- ・都道府県の担当職員に対しては、エラーチェック・訂正の方法について、どのように指導していくのか。本省側の職員の業務体制は十分に確保できるのか。  
→経由機関に極力負担をかけないよう調査手引きの記載の充実や説明会開催の機会を設けることで対応。その際、定員要求といった措置を考えていきたい。
- ・新システムで、都道府県でエラーデータの訂正作業のミスが発生した場合に、問題が起これないような仕組みになっているか。  
→現システムでは、最終保管用は仕掛中、速報用、確報用とステータスを持っているが、新システムでもこれを参考にしながら、ステータスを持たせて上書きしないように対応していく予定。

- ・ 標本設計が変更された場合、新システムで対応可能か。
- 標本設計の変更に応じた対応ができるような設計としたい。
- ・ 仮に非回答が発生した場合、どのように処理するのか。
- 非回答は回答なしとしてエラーが出るようにし、理由が書かれない限りは受理しないこととする。
- ・ エラーチェックのプログラムを職員が修正できるようにするとのことだが、どのように実施するのか。
- システム内にミドルウェアを実装する予定であり、このミドルウェアの設定ファイルの修正に係る知見を有する者を採用することや外部知見の活用を図ることとしている。
- ・ システム全体を運用するための体制はどのように想定しているか。
- 省内に情報システム部門があるので協力いただき、職員がシステムに習熟出来るような体制を整備したい。また、次期システム稼働後の状況を踏まえ、必要な人数を確保したい。
- ・ 新システムの稼働開始までの体制はどのように想定しているか。
- システムの開発とその後の運用にあたり、体制を整えることは大変重要と認識。省内の統計部門は16人しかおらず、体制が弱い。システムをきちんと動かせる人間が事実上ほとんどいないので、定員を確保、充実していきたい。システムの開発にあたっては外部の知見も得て進めていきたい。これまでの状況を踏まえ、システムを改修するために必要な定員や予算確保に向けて、省内で意思疎通してまいりたい。
- ・ 統計部門では専門性、継続性が重要であることから、担当者が比較的短期間で変わるの望ましくないが、文部科学省では、その点どのように配慮されているのか。
- 人事内で統計部門を何度も経験するという運用が行われている。情報システムの専門性の視点は欠けていたと思うので、定員を補って行きたい。一般統計は各部局にまたがっているため、専門性はどうか、人事部局とも相談しているところ。小さな部門だけで回ると外の目が入りづらくなる面もあるので、専門性に十分に配慮しつつ閉鎖的にならないように組織としてマネジメントしていきたい。
- ・ ようやく全体像が見えてきた。3点確認したい。1点目、6ページに相当する見取り図を新システムでも見せてほしい。新システムで達成を目指す柔軟性として、調査項目等の修正における柔軟性と審査段階で経路機関が直接アクセスできるようになることの2点である旨を明記してほしい。2点目は、6種類の調査票のデータがどのような関連づけでシステム内に格納されるのか。現行では、20ページのように個々の調査票が固定された仕組みになっており、柔軟性がないと思うので、その点がどう改善されるのか。3点目は、スケジュール面でどの程度効率的になるのか。見通しについて示してもらえると、新システム導入のメリットが理解できると思う。
- 新システムの見取り図は提示できていないので、今後対応したい。2点目。調査票を審査して、公表する流れは基本的にどの調査でも同じと認識。根幹となるロジックは変わらないと思う。3点目のスケジュールについては、そもそもかなりタイトなの

で、どこまで短縮できるか、見定めていきたい。

- ・ 2点目は調査票ごとにプロセス処理をしているのか。相当に複雑な処理をしているのか。

→現行システムでは20ページのとおり調査票のデータを部分ごとに区切っている。可変長のような分かりにくい構造になっており、分析の妨げになっている。新システムではこれを解消することを目指していく。

- ・ 16ページに「同様のシステムの例が少なく」とあるが、他の統計と業務の流れに違いはないと思う。このシステムの本質的な問題点を解消しないと、過去と同じ認識で同じように作ったら、新システムでもまた同じ問題が再び起きてしまうだろう。20ページにあるような複雑な調査票を見直し、データを構造化して、必要な情報をどのようにデータベースに格納するかをきちんとデータモデリングすることが必要。1:n、2:n、n:nといった構造をたくさん入れるとその後のデータ結合プロセスも含め、システムがどんどん複雑化するし、それを運用の中で人手で補おうとするとミスの要因となることが多い。昭和の時代からの継続性を重んじるばかりに過去の紙の調査票をそのままデータベース化しようとしている点に無理があり、そもそもどのようなデータをどのように収集してアウトプットしたいのか、それに合ったデータを格納するためのデータベース構造はどうあるべきなのか、という検証がしっかりと出来ていないところが、何かを少し変えるだけでも膨大な時間とコストがかかるシステムの複雑さと著しい柔軟性の低下の一番の原因になっているのではないかと思っている。そこに注意されて、新システムを作っていくと良いのではないか。

→長年の経緯で複雑な構造になっている。人間が手作業で行うと間違いが起こるので機械化して間違いが起こらないようにしてきたが、かえってシステムが硬直していて変更ができなくなっていた。同様のシステムについて各省や民間に話を聞いた範囲では例が少なかった。きちんと進めていきたい。

- ・ 3点聞きたい。1点目は、エラーチェックについて、8ページの図を見ると3回行われているように見えるが、3段階で行われているエラーチェックの内容は同じか否か。同じなら、なぜ何度もするのか。違うのであれば、それぞれの視点の違いは何か。また、集計の段階で、インコンプリートなデータはどのように処理しているのか。2点目は、新システムでは、職員が調査項目の変更に応じて、自ら変更できるが、そのことによってどういう効果が期待できるのか、具体例があるのなら教えてほしい。3点目は意見、印象であるが、文部科学省の統計担当者のサイクルは他府省に比べても短いと感じている。学校基本調査は調査事項がほとんど変わっていないのではと思うが、学校教育はものすごい勢いで変わっているのだから、調査項目もどんどん変わって然るべき。統計担当者のサイクルが短いために、現行の仕組みを変えることがさらに難しくなってしまうと感じる。人事の話は統計委員会で議論することではないかもしれないが、学校教育の現場の状況を統計部門でもきちんと吸い上げられるように考えてもらえれば。

→3点目の意見について、調査項目を変えなければならないというのは、同じ認識を持

っている。今回の問題を契機として、局を横断して、統計部門の人事交流を行って相乗効果をもたらすべきであり、方向性は人事課とも共有できている。人事サイクルが短いとの話であったが、同じ人が長く居続けた結果として物事が動かないということもあり、一長一短がある。今後は、状況を見ながら考えていきたい。1点目のエラーチェックはそれぞれ異なるもの。最初の審査の段階で最もエラーが発生する。複数の調査票を足し上げることで、単体の調査票ではわからないエラーが新たに出るので、学校、市町村等が確認の上対応する。2つ目の審査は、個別事情で最初のチェックで対応できなかったエラーについて、理由を確認しながら、本省で個々に対応するもの。3つ目の確報段階での審査は速報後の変更に対して新たに生じたエラーを再度チェックするもの。次に、非回答が発生したときの取扱いは、学校基本調査では、回収率100%を維持、未記入があれば回答できない仕組みとなっている。また、回答した値が異常値であった場合、疑義照会した上で集計している。審査を複数回行うことが、硬直性に跳ね返っているという課題があるが、そのことでデータの正確性を担保しているとも言える。2点目の集計公表作業を分けることの重要性に関しては、2次利用集計するとき抽出・集計機能が重要であるが、自動的に製表されてしまうのがブラックボックスの要因となっているため、一旦、CSVで吐き出す機能を設けることで柔軟性を持たせるもの。公表用の帳票は印刷を重視したつくりになっているため、製表プログラムを用いて対応する。

- これまでの点検検証で明らかになったミスが多くは、印刷や可視性を意識した公表用エクセルに対して調査結果データをコピー&ペーストされる際の手作業ミスだった。わざわざエクセルをセル結合して見栄えを良くする過程で間違い混入のリスクを増やす必要はなく、マシンリーダブルな状態を取り込みができたり、セル結合を外してから加工編集する手間を利用者に強要しない方がデータは活用しやすい。どうしても印刷物が必要な場合のみ対応するとして、基本的にはデータ活用を想定してCSVなど活用しやすいデータで一貫して対応したほうがよい。また、人手を介してデータを加工したり、複数の組織を跨ってデータをやり取りすると、それだけ間違いが発生するリスクが高まるため、調査回答者がデータ入力する時点から正しく入力してもらうために、異常値やヌル値の登録を防ぐための画面の規制や多義性のない選択肢の提示などの工夫がデータ品質向上に有効となるので意識するとよい。
- 現行のシステムも新しいシステムもそうだが、市町村の小学校をイメージされているが、県立高校とか私立学校とか大学など所管が違うところはどうなっているのか。  
→例えば、公私立の高校はまず都道府県でチェック。大学等は本省に直接回答されるので、本省でチェックする。
- 4点ほど。1点目、15ページのスケジュールがタイトに感じる。特に1年目の要件定義。ここでは何を決めて何をアウトプットするのか。これを基に下流工程が決まっていくので、要件定義にかかっている。このシステムは、都道府県や市町村でも使用することになるとのことだが、今の業務をそのまま要件定義に書き込んでいくのか。時間的に厳しいように思うが大丈夫なのか。どうやって乗り切るのか。2点目、デー

データベース化について。調査票の種類が多いが、データベースの構造を作るのは非常に大事なのでそこをどうしていくのか。学校の基本情報をフェイス事項のようにして、そこから生徒数や卒業後の進路といった個々のデータがぶら下がるような構造となるのか。3点目、データベースと関連するが、調査票の扱い方について。将来新しい調査票を追加するときに新しいデータベースを子どものようにぶらさげることに対応できることになるので、モジュール単位で拡張性を持たせてほしい。柔軟性を持たせるときに、データ項目の追加とサブデータベースの追加の2通りあると思うが、どこまで変更できるようにするのか。4点目、エラーチェックについて、処理システムの問題ではないかもしれないが、学校基本調査のかなりの情報は公立学校であれば行政資料で分かるのではないかと。であれば、行政情報のデータベースを整備する方向に注力することも考えられないか。可能性があれば、検討いただきたい。

→1点目、要件定義の日程がタイトであることは我々も認識。仕様書作成に必要なシステム要件をすべて定義する必要がある。今回頂いた新システムのあり方についてのご意見は仕様書に落とし込んでいくが、書き込み過ぎない工夫として、入札時の提案に盛り込んで加点対象とする。都道府県とは、変更点の周知、意見を聞きながら見直していくことも必要と考えている。2点目と3点目のデータベースについては、十分詰めていない。今のつくりはSQL型で相当堅い構造となっている。どういうデータベースを選び、そこに間違いなくマイグレーションしていくことは重要だと思うが、あまりに変えると職員による変更を担保できなくなる可能性があるのでオープンなものも含め十分検討したい。サブデータベースの在り方も業者とも相談しながら検討したい。4点目、公立学校のシステムは非常に複雑。公表時期の兼ね合いや効率的に作業するには、既存の調査の枠組みで情報収集することが現実的。ただし、学校も統廃合を繰り返し、歴史が追えなくなっている課題があり、ユニークな学校コードをつけるべきではという意見もある。その観点から、今後、調査で得るもの、行政情報で得るものを精査し、見直していきたい。

・技術的な話も重要ではあるが、プログラムを複雑化・ブラックボックス化させて変化への対応力を低下させる根本的な原因は指摘したとおり従来の調査票をそのままに踏襲したデータ構造の問題。アーキテクチャを既存のままとすると、基盤をいくら新しい技術に刷新しても既存の問題を引きずってしまうので、見直すなら抜本的に見直した方がよい。そもそもどういうデータを調査しなければいけないのか、どういうデータが動的、静的でそれらが本来どう取り扱われなければならないのかの整理とモデリングについてしっかり委託先事業者と話していただきたい。統計の継続性の確保にも一定の配慮は必要だが、現行の調査票の構造見直しは将来のために必須と考えるので、注意してほしい。

→過去のシステムが持っていた負の部分とは決別したいと考えているので、しっかり検討していきたい。

・宿題は、6ページの見取り図の新システム版を後日示すということだったかと思うが、対応できるか。

→事務局と相談して対応する。

→宿題は後日提示されることになった。また、本日の議論を踏まえた取りまとめは後日の部会で行うこととされた。

## (2) 重点審議（「プログラムミスによる統計数値の誤り発生の防止」の審議の進め方について）

事務局（統計委員会担当室）から、資料2に基づき「委託先におけるプログラムミスによる統計数値の誤り発生の防止」の検討の進め方（案）についての説明が行われ、質疑が行われた。

主な発言は以下のとおり。

- ・別紙の「プログラムミス」の定義が分かりにくい。運用上のミスを含む旨を明確にするなど工夫した方がよい。
  - ・プログラムの設計テストの工程をどのように確認していたのかという観点を入れてはどうか。特に、最終的な統計データのチェックには必ず実査の担当が関わるべき。誰が、どういうチェックをしたのか、プロセスに問題がなかったかを確認したい。
  - ・調査結果をどう使っていくのか。端的に言うと、しっかり発注者が「何をしたいか」という仕様を書きましょう、その仕様どおりにプログラムが適切に動くかをテストしましょう、仕様どおり動いているかを最終的に検品しましょうということに尽きる。また、調査は書面、ワードでやるのか。できたらウェブ画面経由でしっかり回答選択肢の設計を行ってやるべきだが、少なくとも前回同様のワードでやったら回答の多義性や重複が発生するため、やり方を工夫しないと集計が大変。
- 書面調査などでプログラムミスが原因とされたものが結構あって、問題ではないかと委員から指摘いただいた経緯がある。プログラムミスと記載がなくても実はプログラムミスの場合がありうるので、念のため確認するもの。ご指摘のとおり基本的なところだが、それが本当にできているか確認することも大事。今回の不適切な事案は基本ができていなかったところが多々あり、そうであれば調べようというもの。なお、今回の調査はエクセルで行う予定。
- 今回の、特に基幹統計における問題事象を見ていくと、プログラムミスよりも計画変更を公表し忘れていたり、コピー先のセルを間違っ値を貼り付けてしまったといった人為的、運用上のミスがかなり多かった印象である。「プログラムミスによるものだったか」を問う設問があるが、プログラムミスによらない人為的なミスの区分もあった方が数字がはっきり出て価値あるのではないか。
- ・目的意識をはっきりさせることは大事。私なりの頭の整理を申し上げるが、1つ目の大きな目標は、ヒヤリハット事例をきちんと分析して今後の再発防止を考えていくこと。最終的にプログラム委託する際のチェックポイントのリストのようなものを整備できたら、外注する際により適切な管理ができやすくなるのではないか。そのため、

調査はプログラムによるデータ処理に絞った方が回答してもらいやすいのではない  
か。2つ目は、単に「ここに気をつけましょう」だと意味がないので、チェックポ  
イントとあわせて、防止策のグッドプラクティスを探るといこと。3つ目は、エラー  
が起きた際にその組織が本当の原因は何かを追究していくことが大事なので、自ら問  
答を繰り返して原因と再発防止策を考える契機としていただくといこと。この3つ  
を目標としてやっていったらどうか。

もう1点、別紙の「事象の区分」のところ、仕様書ではこう書いたつもりだったが  
受け取った側はその仕様書を誤解したなど、コミュニケーションエラーが分かるよう  
に回答してもらえるとありがたい。

→コミュニケーションエラーは事象というよりも原因に近いので、どこでくみ取るか工  
夫させていただきたい。また、2つ目のご提案の、防止策のグッドプラクティスを探  
るはそのとおりだと考えているが、書面調査よりもヒアリングで対応していくものと  
理解している。

- ・プログラムミスに集中するのはよいが、それであれば、例えば「内部にバグがあるも  
の」という選択肢が選択されて、それが何割あったという結果が出てきても、工程を  
もう少し分解して踏み込んだ設問にしないと、対策に結びつかないのではないか。
- ・省内でプログラムを作ること自体が難しいため専門性がある委託先をお願いしている  
という前提に立つと、職員がどういう点を確認すれば妥当と判断していたかがポイン  
ト。人員体制が頻繁に変わっていく中で、その辺りの確認の仕方をどうしていくかが  
再発防止を考える上で非常に重要と考えているが、そういった点を確認できるような  
質問を設けられないか。

→書面調査の実施時期や回答期間の確保などを考慮する必要があるため、対処方法を検  
討する。

- ・府省へのヒアリングは想定していないのか。

→プログラムミスはきちんとやっていないから起きたものが多いと考えており、ヒアリ  
ングを実施してもあまり審議の参考にならないと考えている。事務局で情報を集めて  
部会に提出し、部会の中ではベストプラクティスを探るほうを主にさせていただければ  
と考えている。

- ・ご意見を踏まえた案はメールで相談させていただき、調査自体を早く進めたい。ヒア  
リングでは業界団体や統計センターからベストプラクティスの取組を聞きたい。

→書面調査の結果を踏まえてヒアリングに臨みたいが可能か。

→書面調査結果を委員やヒアリング先にメールで共有したうえでヒアリングに臨んでい  
ただくスケジュールを考えたい。

### (3) 第1次再発防止策(案)について

事務局(統計委員会担当室)から、資料4に基づき、一般統計書面調査の回答状況  
について説明が行われたのち、資料3に基づき、第1次再発防止案について素案からの  
変更点を中心に説明が行われ、意見がある場合にはメールで事務局に連絡することとさ



れた。

(4) その他

次回部会日程は6月20日(木)、場所等詳細は改めて連絡する旨、事務局から案内された。

以上

<文責 総務省統計委員会担当室 速報のため事後修正の可能性あり>