

統計データ利活用に向けた国・都道府県の協働について

総務省統計局・（独）統計センター 統計データ利活用センター長 谷道 正太郎

1. 統計データ利活用の動きと『統計データ利活用センター』

現在、政府では統計改革の大きな動きが進められている。2017年5月には、統計改革推進会議（議長：菅義偉官房長官）が「最終取りまとめ」を決定するとともに、2018年3月には政府全体の統計整備に関する総合的な計画（「公的統計の整備に関する基本的な計画」）が閣議決定され、また、5月末には統計法が改正されるなど、まさに10年に一度の大改革が行われていると言える。

今回の改革の大きな特徴の一つとして、「統計の改善・精度向上」の面だけではなく、統計データ等の客観的証拠（エビデンス）に基づく政策立案、いわゆる「EBPM（Evidence-based Policy Making）」についても、車の両輪として推進していくことが挙げられる。つまり、統計データ等の質の向上により、より良いエビデンスを基にEBPMが推進されるとともに、EBPMの取組を進めることで、必要な統計データ等についてのニーズ・要望が顕在化し、それが更なる統計データ等の改善につながる、という好循環を生み出すことを目指している。また、そのための環境・基盤の整備として、社会全体における統計データ利活用を促進していくことなどが必要とされている。

こうした中、2018年4月に、総務省統計局と独立行政法人統計センターは、価値創造につながるデータの利活用が広がるよう様々な取組を進める“先進的なデータ利活用の推進拠点”として『統計データ利活用センター』を和歌山県に開設した¹。

本稿では、当センターの取組内容とともに、都道府県との協働による統計データ利活用の姿について紹介することとしたい。

2. 統計データ利活用センターの取組の概要

まず、統計データ利活用センターの業務概要であるが、当センターのコンセプトを一言で表せば「先進的なデータ利活用の推進拠点」となる。そのために、業務として、以下の三本柱の取組を進めることとしている。

一つ目の柱は、高度な分析を実現するための「統計マイクロデータの提供」である。

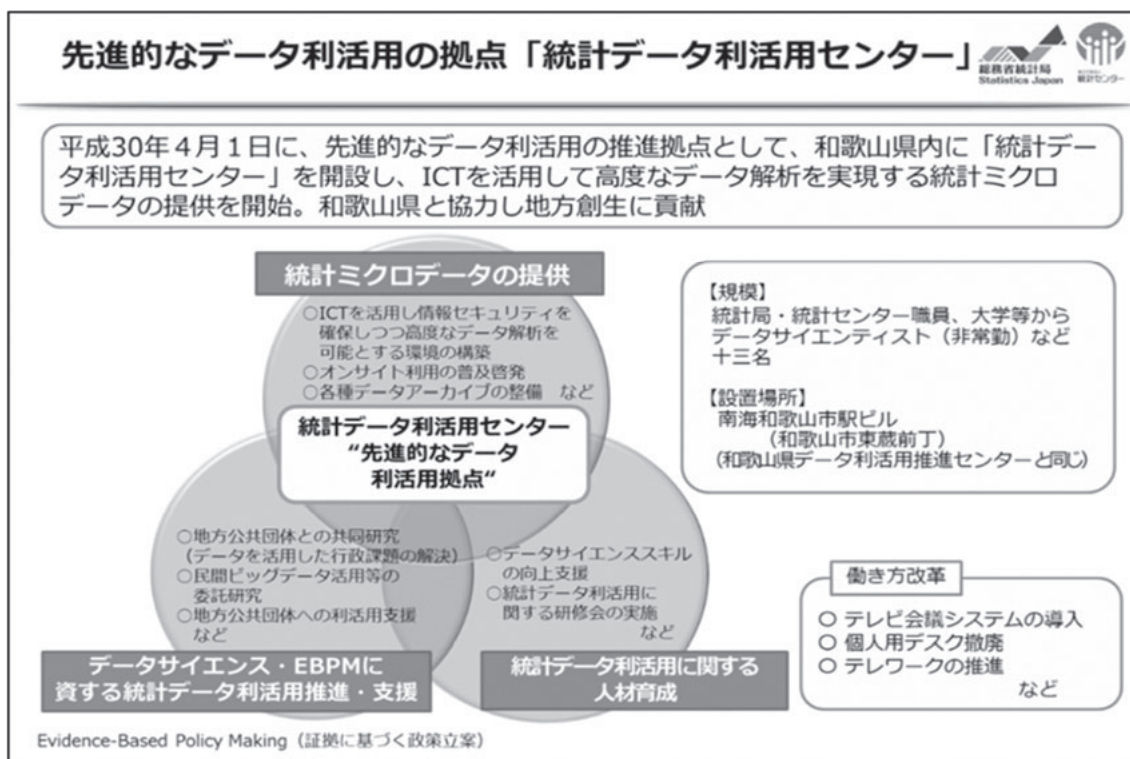
¹ 当センターの和歌山開設のきっかけは、政府が進めている地方創生の取組の一環である「政府関係機関の地方移転」の動きによるものである。和歌山県からの誘致提案を受けて検討がスタートし、検討の結果、地域の課題解決や発展に資する統計データ利活用を、和歌山県をはじめとする各県との協力や産官学との連携により進めていく拠点を和歌山に設置することとなった。

情報セキュリティを確保した分析環境（オンサイト施設）を構築し、統計マイクロデータを提供するとともに、オンサイト利用の全国展開を推進している。

これにより、信頼性の高い公的統計データがより多様に活用され、社会や経済の発展を生み出す新たな発見につながっていくことが期待される。また、この取組を発展させていき、行政が保有する各種情報や民間ビッグデータなど、幅広いデータを用いて様々な分析ができる環境の構築や展開ができないかなど、ビッグデータ時代をリードする拠点としての可能性を様々に追求している。

二つ目は、「データサイエンスやEBPMに資するデータ利活用の推進・支援」である。地方公共団体においても統計マイクロデータを始めとする各種のデータを活用した課題解決の取組は重要となっている。これに対する支援や共同研究を実施するとともに、優良事例を広く展開している。

三つ目は、「データ利活用に関する人材育成」である。データから価値を生み出すことのできるデータサイエンティストの育成は、今日の大きな課題となっている。こうしたことから、公務員向けの研修会や、広く社会の統計リテラシーの向上にむけたセミナーの開催などの取組を進めている。



出所：総務省統計局・独立行政法人統計センター説明資料

当センターでは、これらの活動を通じ、我が国における社会経済の発展や、地域の課題解決に資するデータ利活用の姿を全国に発信している。

また、当センターの開設とともに、和歌山県では「和歌山県データ利活用推進センター」が開設された。両者は同じ場所に併設され、連携・協力しながら、データ利活用の良いモデルの構築を目指している。

次に、これらの取組をより詳細に見ていきたい。

3. 統計マイクロデータの提供

（1）統計マイクロデータによって広がる可能性

統計調査により集められたデータは、集計され、クロス表などの形として公表されている。しかしながら、より高度な分析、多様な観点からの分析を行おうと思うと、集計される前のデータを用いることが必要になる。

「統計マイクロデータ」（「調査票情報」）とは、集計される前のデータのことであるが、これは、行政機関などが統計法に基づき一定の公益性を有する統計作成や分析を行う場合に利用することができる。統計マイクロデータを用いることで、上述のように、公表されたクロス表以外の自由な観点での集計や、表作成以外の様々な手法による分析などを行うことが可能となる。

これまでも、地方公共団体において、統計マイクロデータを用いて、学区別による集計や、地方における消費者物価指数など独自指標の作成、産業連関表の作成、各種計画策定のための基礎資料作成などが行われてきているが、最近では、ICTの進展によって統計分析ソフトも充実し、機械学習をはじめとする様々な手法による分析もより身近なものとなったことから、統計マイクロデータの持つ可能性もますます大きなものとなっている。また、EBPMの必要性の高まりに対応し、統計マイクロデータの更なる有効活用を図ることが求められている。こうした中で、統計マイクロデータの利用における情報安全性の確保も、当然ながら徹底される必要がある。

（2）データのオンサイト利用の試み

こうした背景を踏まえ、現在、統計マイクロデータの利用をより安全に、そしてより自由な形で分析できるための新たな環境整備が試行されている。

これまでの統計マイクロデータの利用においては、①統計マイクロデータ所管省庁への利用申請の段階で詳細な分析内容を明らかにすること（例：統計調査のどの調査事項のデータを用いて、どのような分析手法・集計内容で、どういった結果を作成するのか等）が求められ、その上で、②当該分析のために必要となる最小限の調査事項のデ

ータだけが提供される、という方法がとられていた。そのため、利用者は、①事前審査のプロセスに大変時間をとられたほか、②実際に手元に届くデータも、限られた最小限の調査事項のものだけにとどまったことから、あらかじめ申請していた決まりきった分析しか行うことができず、試行錯誤を繰り返しながら行う探索的・創造的な分析をすることは困難であった。

上記のように、これまでデータ提供の方法がかなり制限的であったのは、データがDVD等によって提供される（利用者側にデータが媒体ごと渡ってしまう）ため、情報セキュリティの確保が難しかったことが理由の一つとして挙げられる。

今般、新たな提供方法として試行されている方法は、情報セキュリティを確保しつつ高度なデータ解析を可能とする環境（オンサイト施設）を整備し、利用者はオンサイト施設内で統計マイクロデータを用いた分析を行う、というものである。この提供方法であれば、国側がデータ利用のセキュリティを確保する（利用者側は外部にデータを自由に持ち出せない²）こととなるため、利用者としては、利用申請の段階では分析の概要だけを申請すれば足りることとなる。こうしたことから、利用者側では、①従来に比べ事前審査の時間が大幅に短縮されるほか、②多くの調査事項のデータを利用することができるため、より自由度が高く、探索的・創造的な分析を行うことが可能となる。

オンサイト利用においては、外部のデータの持ち込みも可能になっている。そのため、例えば、自治体が保有する行政情報や、自治体独自の調査データと国の統計マイクロデータを組み合わせた分析なども可能となる。現時点³でオンサイト利用可能な統計マイクロデータは、総務省統計局所管の統計調査と、経済産業省所管の基幹統計調査のデータとなっている。今後、各府省の統計マイクロデータについても、オンサイトにおいて利用可能となることが期待されており、段階的な拡充を図ることとされている。



オンサイト施設における分析の様子

また、オンサイト施設は現時点で9か所において試行的に運用されており、改正統計法の施行（2019年5月1日予定）にあわせた本格運用を目指している。今後、オン

² 分析結果をオンサイトから外に持ち出す際には、匿名性が担保されているか等の情報安全性の審査を受けることになり、この審査を通過した分析結果のみがオンサイト外に出すことが可能（公表可能）なものとなる。

³ 2019年1月時点。

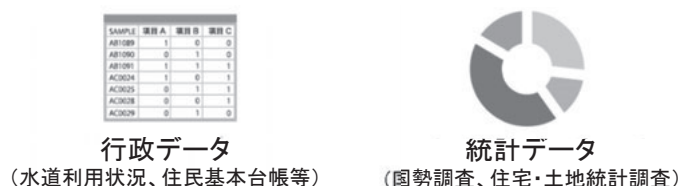
サイト施設の全国展開も視野に入れられており、全国各地において、マイクロデータを利用した新たな発見が生まれることが期待されている。

4. データサイエンスやEBPMに資するデータ利活用

統計データ利活用センターでは地方公共団体に対し、統計データを活用して地域の課題解決を支援する取組を行っている。

データは、それ単体で用いるだけでなく、他のデータと組み合わせることにより、得られる情報は大きく広がる。統計改革推進会議「最終取りまとめ」においても、「統計及び統計マイクロデータの更なる利活用とともに、新たに行政記録情報や地方自治体・民間が保有する各種データの積極的な利活用も統計システムに組み込んで、統計等データを始めとする各種データを有機的・効果的に利活用した統計的分析などを積極的に促進する」とされている。

ここでは例として、「空き家対策」をテーマとした共同研究を紹介する。本件は、和歌山県・和歌山市とともに、東京大学空間情報科学研究センター・助教秋山祐樹氏の研究室より協力を得て取り組んでいるものであるが、国勢調査や住宅・土地統計調査といった統計調査のマイクロデータとともに、行政データ（例えば、住民基本台帳や水道の開栓情報など）を組み合わせ活用し、空き家に関する分布の推定や将来推計を行うことを目標としている。また、これらの分析の実施においては、前述のオンライン環境も利用しながら進めていくこととしている。



例) 人口・住居データ×水道使用データ
⇒空き家の分布を推定

出所：筆者作成

具体的な取組としては、「①空き家分布推定手法の構築」として、行政データ及び統計データを活用し、空き家が数多く分布している地域を推定・把握できる手法、推定モデルの構築を研究している。これにより、新たに空き家調査を行うことなく、既存の情報（行政データや統計データ）から迅速・安価に空き家の現状把握・推定を行うことが可能となることが期待できる。

また、「②将来の空き家分布推定の研究」として、空き家分布推定の手法を活用し、将来空き家が増える可能性が高い地域を推定・把握する手法を研究している。将来予測ができれば、優先的に対策を講じる地域の絞込み等を行うことが可能となるなど、施策への応用範囲も広がることが期待できる。

そして、「③推定モデルの社会実装・展開」として、空き家分布推定の手法・構築モデルを活用し、各地の空き家の把握に応用することや、空き家把握以外への応用・展開について研究することとしている。

5. 人材育成

データ利活用の推進においては、データを実際に扱うスキルや、データを正しく読み解くリテラシーなど、人材育成の面も非常に重要になる。

こうしたことから、当センターでは、統計マイクロデータを用いた分析手法や、データを地図情報と組み合わせて可視化する手法（統計 GIS の活用）などに関するセミナーを開催するとともに、都道府県において開催される研修への講師派遣なども行っている。こうした研修における連携をきっかけとして、4. で紹介したデータ利活用に関する連携につながることも多い。

6. データ利活用・EBPMを進めるために

最後に、データ利活用・EBPM の実践を進めるためのポイントについて述べたい。

当センターでは、2018年6月に、全国都道府県の統計部局の利活用担当者が一堂に会し、統計データ利活用を推進するための会議を行ったが、その中の「データ利活用」にかかる意見交換において、多くの担当者が共通して感じている主な課題は「①スキル・人材育成」、「②組織文化・体制」、「③データ面」の三点であった。

このうち、「①スキル・人材育成」に関しては、データ利活用のスキルを有する職員が少ないことや、そのために利活用の取組が属人的なものになりがちであること、また、人事異動サイクルが短いため育成が難しいといったことが挙げられた。また、「②組織文化・体制」における課題とは、利活用を進める土壌が組織内にないといったことや、部局間の役割分担が不十分であることなどであった。そして、「③データ面」については、市町村別のデータが整備されていない場合があることや、データを

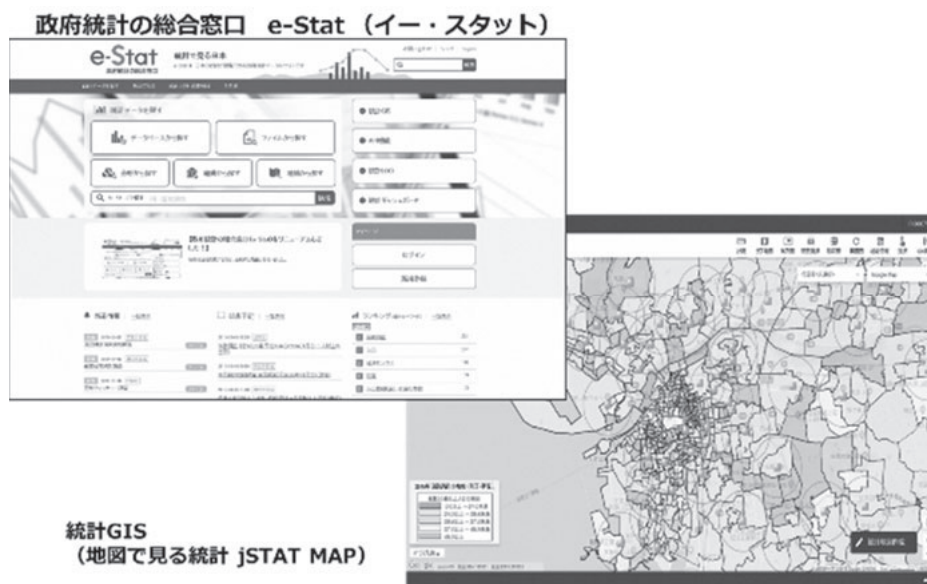
見つけることが難しい、といったものが挙げられた。

これらの課題を克服していくためのキーワードの一つとしては、「連携」が挙げられよう。

例えば、「①スキル・人材育成」に関しては、地元大学等との連携により、アカデミアの有する知見を活用するとともに、役所側の異動サイクルにとらわれない長期的・継続的な取組が期待できる。「②組織文化・体制」については、まさに組織間の連携を図ることが重要であり、その中で好事例を一つでも生み出し、共有していくことが求められる。「③データ面」についても、マイクロデータの活用のほか、欲しいデータが直接にはなくとも、それを代理するデータや関連するデータがないか、さまざまなデータ・ホルダーとの連携のなかで探しつつ、考えていくことが必要である。

データ利活用・EBPM の取組は、簡単に結果が出るものではなく、試行錯誤を繰り返しながら、ブラッシュアップしながら進めていくものである。そのプロセスにおいては、様々な方からの知見や専門性を活かしていくことが重要である。

当センターとしても、各地方公共団体をはじめ、様々な方と連携・協力して良いネットワークを構築し、データ利活用の取組をひろげていきたい。



出所：政府統計の総合窓口（e-Stat）及び統計 GIS により筆者作成