

第6回 統計技術・データソースの多様化等検討会

(第15回 ビッグデータ等の利活用推進に関する産官学協議のための連携会議と同時開催)

議事概要

(開催要領)

日時：令和3年11月17日（水）14:00～16:00

場所：総務省第二庁舎6階特別会議室（Web会議併用）

(議事次第)

- 1 開会
- 2 議事

- (1) クラウド型会計ソフトから得られる決算データについて（株式会社TKCの取組）
- (2) EBPM推進委員会データ利活用WG取りまとめについて
- (3) 第5回観測データ利活用検証WGの報告
- (4) 取りまとめに向けた論点について

- 3 閉会

(配布資料)

- | | |
|-------|-------------------------------------------------------------------------|
| 資料1 | クラウド型会計ソフトから得られる決算データについて（TKC） |
| 資料2 | EBPM推進委員会データ利活用WG取りまとめについて |
| 資料3 | 第5回観測データ利活用検証WGの報告 |
| 資料4 | 取りまとめに向けた論点について |
| 参考資料1 | 第14回ビッグデータ等の利活用推進に関する産官学協議のための連携会議
（第5回統計技術・データソースの多様化等検討会 同時開催）議事概要 |
| 参考資料2 | 統計技術・データソースの多様化等検討会の開催について |
-

(概要)

【(1) クラウド型会計ソフトから得られる決算データについて（株式会社TKCの取組）】

株式会社TKCから、資料1に沿って説明。主な発言は以下のとおり。

- 最近自分も企業の財務諸表のデータを解析しているが、他の企業だと、例えば生産性の高い企業にデータが偏っているという非ランダム性のサンプリングの問題がある。御社の事例でいうと、全体の10分の1がカバーされているということで、ランダムサンプリングで10分の1ということであればかなり精度が高いと思うが、この部分はどのように考えているのか。地域集計の際などに、補正などはしているのか。
 - どうしても異常値は出てくる。本日ご紹介したのは月次BASTというものだが、既に年度版BASTというものも出しており、多くの金融機関や大学などで使っている。そこでは異常値は排除して集計しているという実績もあり、その基準は月次BASTでも採用している。

- 会計士の先生から実際に集めたデータということで正確性も高く、月次なのでコロナの影響も反映しやすい、とてもいいデータだと思うが、無償でこういった素晴らしいサービスを提供している狙いは何か。また、25万社と聞くと、もっと御社のユーザーは多いのではないかという気もするが、この数字の根拠は何か。また、これだけのデータがあれば、他の会計ソフトと組めば色鮮やかに日本全体を映し出すことができるのではないかと思うが、そういったことは御社にとってメリットとなるか。
 - 無償提供については、コロナ禍という状況もあり、足下の景況感をつかむにはどうしたらいいのかという声を直接いただいており、公共性の高いデータだということで、無償で出すことを決めた。先ほどご紹介した年度版BASTは、覚書を締結している金融機関などに無償提供しており、顧客支援の中で使ってもらおうということで、今後も価格をつける予定はない。25万社というのは、先ほども異常値という話があったが、巡回監査等できちんと精査したデータでないと提供できない。TKC会員（税理士・公認会計士）の顧客のデータになるので、精査した結果としてこの数になっている。

- 計上されていない会社もあるということか。
 - 了承をいただいたTKC会員のデータになる。

- 個人事業主が入るかどうかが活用の仕方が変わるかと思うが、個人事業主は含まれるのか。
 - 含まれていない。

- 金融機関にも自動で報告できるという話について、政府でも統計の報告者負担の軽減のために民間の会計ソフトと連携するという話が何年か前にあったが、その話とも親和性が高いと思う。御社として、統計調査への報告と連携することはありえるのか。
 - 金融機関にも自動で報告できるサービスとは、「TKCモニタリング情報サービス」

(TKC会員が毎月の巡回監査を実施したうえで作成した月次試算表、年度決算書などの財務情報を、関与先企業の経営者からの依頼に基づいて、金融機関に開示する無償のクラウドサービス) のことで、TKC会員が税務署へ電子申告した同じ決算書データを、金融機関に送ることで、実際に資金繰りに苦しんでいる企業に対して早急な支援ができるのではないかという思いから、5年前から行っている。特許も取っている仕組みなので、国の統計調査に役立つものであれば良い話だと思う。経営指標については、一般的には年度版のデータが多く、月次版のデータは世界にもあまりないのではないかと思う。更に改良していきたい。

- このデータは公的な活用への価値が高い。例えば有償化して、一部暫定版として速報を出すなど、カスタマイズしたサービスを出すという見通しや要望はあるのか。
→ 今のところ、そういった要望は聞いていない。基本は会員事務所の監査が終わったデータということで、精度を大事にしているので現在はこのようなサービスの公開としている。
- データとしては残高試算表の係数を取っているということか。そうだとすると、売上や利益以外に固定資本などのデータも取っているのか。また、タイムラグについて、2か月と25日あるということだが、これをあと10日早めると大分データ数が減ってしまうのか。
→ データについては、1ヶ月と25日かかる(9月末時点のデータは11月25日で締めて、30に公開)。実際、会員事務所では月末に監査をするところも多いと思う。ボリューム感なども踏まえると、どうしてもタイムラグがある。基本的に送っていた会計データなので、固定資本などのデータもある。こういう分析値がほしいという要望があれば、開発に着手することも考えられる。

【(2) EBPM 推進委員会データ利活用 WG 取りまとめについて】

農林水産省木村統計企画管理官から、資料2に沿って説明。主な発言は以下のとおり。

- 民間と公的統計、それぞれの長所を生かして補うというのはとてもいい考え方。その際、補正をどういった形で行うのが技術的に一番難しいポイントだと思っている。例えば農産物の流通・消費について、例えばPOSデータとトラックのGPSデータの関係が分かると、お互い別のデータで補正する、といった事例が積み上がっていくと思う。また、民間からデータを借りる場合、目的をはっきりさせて信頼を得た上で借りるという話だが、目的をどれくらいはっきりさせるかは難しい。例えば分析や研究の場合、もっと違うことにも使えるという方向に発展すると、当初の目的とはかなり変化すると思う。目的をどの程度の粒度で約束するかという議論はあ

ったのか。

→ 農水省の例だと、価格を調査する場合、例えばこれまでは生産者や卸売業者などへ調査するという観点しかなかったが、世の中にあるデータをどのように活用していくか、我々もいろいろなご知見をいただきながら考えたい。

また、「行政と民間の間で合意すべき事項」の中の、「データ利用目的の粒度」については、ケースバイケースで、相手との関係の中で決められていくものと考えている。ただ一般論としては、やはり株主への責任もあるので、データの使用範囲などを出来るだけ限定する形が望ましい、提供する民間企業としてはそれが原則論だ、という意見もあり、このような文章になっている。

● EBPMの基盤になるデータということは、将来的に検証可能である必要がある。研究の場合も、説明可能性のために10年はデータを保存するといったことがされるが、民間企業が同じデータを保持し続けるのかということがある。また、きれいに処理された公的統計のデータとは違い、エラーの処理などもしなければならぬので、修正した際のソースコードの公開や保持が必要かと思う。検証を可能にするためのデータの取扱いなどはWGで議論されたのか。

→ 「合意すべき事項」の中で、個別の事例に応じて盛り込みを検討する事項の中に、提供されたデータの保有期間、保有期間終了後のデータの取扱いについて言及がある。複数の事例と、こうした考え方がEUでも大分整理されてきているという有識者のご指摘に基づき、EUのドキュメントも参考にしながら記載したもの。

● 官民で協力してデータを政策に活かそうということであれば、プロセスの検証可能性もさることながら、使ったデータやその結果、効果について、官民両方で説明責任を果たして検証可能にしていく必要がある。オープンデータにするのが望ましいが、そういうことができるものばかりではないと思うので、少なくとも限定公開、ロジックの公開などはできればと思う。

● 官民がパートナーシップを持つには、官側からこのためにこの統計をこうしたいというのがあって、それを実現できる民間データはどのようなものがあるか、ということが起点になる。民間はどういうデータを持っているかと言われても、民間は基本的に業務を回すため、ビジネスをやるためにデータを生成しているので、その目的にあったデータしか持っていない。こういう目的・性質のデータがほしい、なぜならばこの統計を自動化するため、などがあって、それに適する民間データがどこにあるかという順番に持っていく必要がある。農水省でのPOSデータを使っている事例はとてもいいやり方だが、生鮮品はJANコードがなく、うまくデータを出すのは難しい。また、そのデータは経産省や総務省の統計などでも使えると思うので、

そこを皮切りに、一つのデータを他の統計にも使えるような仕掛けを作るといいと思う。

→ 農水省の調査研究も率直に言って難しい論点はある、ご指摘のようにPOSデータがすぐには調査票の代替にはならないと考えているが、勉強はしていかなければならない。また、まさにおっしゃるとおり、対象者の負担軽減も図りたい、ということもこの調査研究を始めた大きな動機の一つ。先ほどの株式会社TKCのご説明の関連で言うと、農家の方が税理士などに経営データを渡していれば、それを活用できるのではないかと勉強も始めているところ。それぞれの省庁でもこういう勉強はしていると思うが、農水省として調査研究などで得た成果は政府内でしっかり情報共有していきたい。

● パートナリシップの話だが、統計当局の役割は、自分たちで統計を作成して公開するだけではなく、社会全体としてデータが公的な利益のために使われるような状況を作ることもあるのではないか。データを公開するのは事業者か統計当局か分からないが、事業者が統計を作成するのをサポートするという考えられる。事業者は自分で統計を出したくない、事業者しかデータを持っていないので検証しにくいなど、状況によるだろうが、いろんな形があるかと思う。

POSデータについては、バーコードは生鮮食品についていないことがあるので難易度が高い。事業者にも役に立つ仕掛けが何か提案できるとよい。

【(3) 第5回観測データ利活用検証WGの報告】

ビッグデータ等の利活用推進に関する産官学協議のための連携会議（以下「ビッグデータ連携会議」という。）水野構成員から、資料3に沿って説明。

【(4) 取りまとめに向けた論点について】

ビッグデータ連携会議事務局から、資料4に沿って説明。主な発言は以下のとおり。

● ビッグデータが有する偏り、特徴・特性について、メタデータを集めるというのは非常に有効。その際に気になるのは、どこかが責任を持って集中的に集めるのか、それとも単にその都度確認するだけなのかということ。メタデータの管理の仕方について、どう考えるか。

→ おっしゃるとおり、一過性のものだと広がりが無い。せつかくメタデータをただけなのであれば、広く政府の中で共有するところがあって然るべきかと思う。まずは個別事例を整理しつつ、ゆくゆくは共有していくといったイメージで進めていきたい。

- メタデータの例にも関連するが、民間企業の中でもまだ整理できていないところが多々ある。ビジネスをやるためにシステムを作っているのに、同じようなメタデータでもあちこちで別の名前になっているということが往々にしてあり、当然企業によってメタデータも違う。決算データ、納税データ、POSデータなど、ある程度のかたまりでマッピングしたものを出してもらうなどでないと、かなり大変になるかと思うので、その辺りは工夫した方がよい。また、この論点はEBPM推進委員会のデータ利活用WGの成果とも通じる話であるが、どう体系化されていくのか。実効性のあるものにするためにどうしていく予定なのか。
- 既にWG等で得られた知見は盛り込みつつ、公的統計への利活用に特化した部分もあると思うので、より専門的な内容も入れながら進めていきたい。

- これまで連携会議で扱った事例には、現在の公的統計を置き換えたりサポートしたりするものと、新しい公的統計を作ろうとするものがあった。既存の公的統計を置き換えるものと新たにビッグデータの特徴を活かしたのものを作るという2つの系統があることを示すと特徴が出るかと思う。
- 8月の統計委員会企画部会でも、委員からまさに御指摘のような問題意識を頂戴した。改めてそのような視点を持って整理していきたい。

- メタデータの重要性には同意。また、データが得られなくなるところでバイアスが生まれることもあるので、欠測データの発生状況などもデータの範囲に含めて注意する必要がある。データの継続性については、もらったデータを漏洩しない、プライバシーを守るなどのセキュリティ面は大前提として必要。また、インセンティブもかなり難しい問題だが、ひとつはCSR、ESGのSのところにも位置付けやすくするなどもあり得るかと思う。

- 人材育成の話も重要。人によって様々なバックグラウンドがあり、ドメイン、数学、ITなどいろいろなタイプの人がいる中でどのように進めていくか。
- ビッグデータを取り扱っている部署自体がまだ少なく、政府の中での人事運用の問題もある。将来的な課題として、個人のスキルとミッションの中での役割等、より具体的に考えていきたいと思っている。

- 民間ビッグデータ活用等の新たな取り組みや、時代に合わせた統計にモデルチェンジしていくためには、これまでの慣習的に行われているような、（調査開始当時は有益だったが）今の時代において役割が低下している統計も存在すると思われる、利用頻度や統計の目的に照らして棚卸しを行い、必要に応じて停止すること

で、統計に関連する職員の方々が活動しやすくすることも重要ではないか。

(以 上)