

令和4年度経済産業省統計調査名簿整備・活用アドバイザーグループ会議

議事要旨

1. 会議開催日程

第1回 令和4年11月4日（金）10:00～12:00

第2回 令和5年1月23日（月）14:00～16:00

計2回開催

2. 検討事項（論点）

「経済産業省統計調査名簿整備・各種レジストリーとの連携に係る調査研究」事業※に対し、以下（1）～（3）の論点について、制度、技術、運用に係る専門的な観点からの検討を行う。

※「経済産業省統計調査名簿整備・各種レジストリーとの連携に係る調査研究」事業：統計データの利便性を改善し利活用の向上を図るとともに、統計調査業務を効率化し調査結果精度の維持向上を実現させることを目的に、経済産業省が実施する各種統計調査の名簿情報のクレンジングや、法人番号、共通キー、緯度経度情報等の付与、他のデータベースとの連携可能性について調査を行う事業。

（1）事業所レベルの共通キー付与体制

- ・事業所レベルの共通キーの付与体制について、どのような形であれば、特定の事業者、製品、技術等に依存することなく他社に引き継ぐことが可能となるか。
- ・調査票情報の利活用の観点から、企業の企業名変更・合併時の共通キーの付番は、どのような形が望ましいか。
- ・統計調査毎に保有・管理している名簿情報について、データ共通化・正規化の観点から、どのように保持すべきか。

（2）名簿情報・共通キーの提供のあり方

- ・名簿情報・共通キーを提供する対象者はどの範囲とすべきか。
- ・提供するデータの範囲はどのように設定すべきか。
- ・どの程度のマッチング精度の共通キーを提供すべきか。マッチング精度に関する情報も併せて提供すべきか。
- ・統計調査にない項目（緯度経度情報等）やクレンジング済み名簿情報の活用について。

（3）他のデータベースとの連携のあり方

- ・他のデータベースと連携に向けて、どのようなことに留意する必要があるか。

3. 委員のご発言要旨

●事業所レベルの共通キー付与体制について

- ・共通キーの番号体系（ID）はランダムであることが望ましい。背後でマスタを作成し、ランダムな ID と各情報が紐付く形が良い。ID に都道府県番号や市区町村番号を設定してしまうと、変更時に管理が難しくなる。
- ・ID はランダムの方が絶対良い。これから考えるのであればランダムの方がすっきりしている。
- ・ランダムが良い。その場合、属性を紐付けないと意味がなくなるので、その属性をどうするかが論点になる。

●名簿情報・共通キーの提供のあり方について

- ・作成されたコンバータは、二次利用申請者全員に配布するのではなく希望者のみに配布すると利用希望者と活用状況が確認しやすい。例えばコンバータを機密データと位置づけて誓約書を書いてもらい、利用後は削除すること、他人に渡さないことを守ってもらった上で活用すると活用状況を捕捉できる。これは、コンバータが機密性の高い情報を含む場合に、その情報を守ることに繋がる。またコンバータが作成途中で改善の余地がある場合にもそれを理解した人だけが使用するという点でも良い。コンバータが完全に完成し、機密性が低い状態で提供できるようになったら、HP 上に公開してダウンロード可にすれば良い。
- ・研究利用に加えて政策立案や評価に活用することを考えると、緯度経度の位置情報が紐付いているのは素晴らしい。これにより、移動や人流、地理情報との接続が可能になる。しかし、一方で緯度経度情報は高度な機密情報であるので、利用者の範囲や取り扱いには取り決めや注意が必要である。
- ・コンバータには産業に関する情報も付与されると良い。各調査の分類コードに加えて、共通の産業コードを時系列の変遷も含めて附帯させることで、EBPM 評価の際などに業種ごとの分析や集計が可能になる。
- ・緯度経度情報は、表記揺れ等の問題があり管理が難しい住所情報よりも、たいへん価値がある。ピンを敷地のどこに打つかもしっかりと検討する必要がある。
- ・共通キーの提供対象者の範囲を所属組織で判断するのは良くないと思っていた。提供範囲を少し広めに取りたいのであれば、リテラシーや目的で判断するのが良いのではないか。
- ・提供情報について、これまで最新データが見られれば良いという感じだったのが、時系列で見られた方が良いという意見が国際的にも出だした。他省庁でもこのような議論がされており、余裕があれば他省庁の取り組みも確認してみるとよい。

●他のデータベースとの連携のあり方

- ・コンバータの ID 番号には意味のある情報や個票の情報は一切入らず（ランダム）、調査時点の事業所（企業）番号とのみ接続されている状態、つまりは個票を提供された人に

とってしか意味がない情報（表）とすることで、名簿情報や個票情報を守ることができる。

- ・今回作成する共通キーやコンバータが、事業所・企業の代表的なコンバータと認知され、他府省庁の統計調査や民間のビッグデータも接続したいという要望が出て、時間を経て成長していけば良い。そのためにも緯度経度、証券コード、産業コードなど、他の情報を接続しやすいものは附帯することが望ましい。その場合でも附帯情報は背後のマスタで管理し、利用者の目的によって必要な情報を附帯する方法を取るべきである。特に事業所や企業の位置が特定され、機密性の高い緯度経度情報の提供の範囲と取り扱いには慎重を期すことが望ましい。
- ・個人的な意見として、今の時点では他組織の名簿やコンバータ作成の動向に過度にとらわれず、自分たちに必要なこと、やりたいことを実現することが大事。
- ・たとえば財務情報を例にとると、EDINET と gBizINFO は、レジストリとカタログ、サマリー情報が gBizINFO、金融商品取引法の開示業務の原本である財務情報があるのが EDINET という関係。そのような形で利用体系を見ていけば良い。属性自体もカタログレベルの属性、サマリーレベルの情報と実際の値の集まり全体みたいな形で議論していくと少し進むのではないかな。いきなりデータ自体を共有するというよりは、業務的な各項目の意味自体を共有する仕掛けを政府で持ちカタログを運用し、レジストリで情報のありかを発見できるようにする、その辺りが入り口にならないかと思う。
- ・将来に向けては、人に関するデータベースでは読み仮名の付与が問題になってきている。ID が振られていないものと統合する時に企業情報の読み仮名をどう処理するか問題意識を持っている。海外のデータを意識すると英語表記の扱いも課題になる。連携については他省庁の取り組みとあまりに違うと後で大変なことになるが、こちらはこちらでどんどん成果を一つずつ出していくことも大事。

●ベンダーロックインについて

- ・事業者変更に備え、ブラックボックス化を防ぐため、職員への知識共有と簡潔なマニュアルを作成すると良い。事業者とは事後も質問・情報共有し合える関係を構築できると良い。
- ・自身が持つ情報資産を特定ベンダーのソフトでしか動かないものを作らせてしまうことがベンダーロックイン。ID の仕様や紐付ける属性を明確にしていけば、理論上誰でもできる作業となり公正な入札を経た事業とすることができる。ただし、事業者が持つ属性のデータベースの質によってマッチング率が変わるため、そこは評価対象にする必要がある。
- ・ロックインが起こることにはある種合理性があるのだろうと思う。ロックイン解除により失われるものもあり難しいことと思う。重要なのは透明性。暗黙知を説明可能にする等、他の事業者に継承されるように透明性を持って記録することが大切。

●ヒアリング結果について

- ・統計調査担当課室からの貴重な意見であり、なかなか生の声で聞ける機会が少ない情報である。総務省統計局の皆さんにも共有し、経済産業省の事業所母集団 DB の利用の実態や考えを共有できると、この事業を越えたところでも意味がある。
- ・事業所母集団 DB との関係性や各統計調査名簿の一元管理については、デジタル化が進んでいる現在、情報を分散保有している現状のまま、うまく取り寄せて使える様にすることが考えられている。今回のケースでは、まず事業所を指す一つの ID を決めて、それに紐付けた住所や電話番号等の最小の属性情報を共有することで、各統計調査のデータベースに手を付けずに情報をそれぞれ取り寄せて、統合データベースを作った時と同じ情報が得られるモデルが有効に機能するのではないかと。
- ・随所に今の業務で手一杯という意見がみられた。全体的にプロセスを見直して業務が楽になるように、プロセス全体を改善する方向でも見ていった方が、最終的には目的を達成できるのではないかと。

●マッチング結果について

- ・マッチングの結果が、調査によりかなり違いがある。記述の仕方が違う点は揃えれば揃えられる。調査の性質上、避けられないものと改善で解決できるものがあると思う。
- ・調査票への記述の仕方が違うという点では、レンタルオフィスを活用する会社が増加するなど、実態が変化する時に統計の枠が変化する、というような話とも関係しているのかなと思う。データそのものばかりでなく、データ項目や実態を取るためのデータ構造にかなり様々なことを考えてやられていると思うので、その辺りが見えてくると面白い。
- ・共通キーを付与した成果として、共通キーの付与率に加えて、共通キーを付与したことにより各統計調査同士がどれくらい突合できた（マッチできた）かが分かると良い。

●METI 共通キー等の提供ルール案について

- ・金融庁の EDINET では、法人番号、EDINET コード、証券コード、法人名から検索が可能であり、法人番号や EDINET コード、証券コードの内、どれか1個くらいあれば足りるのではないかと。他が管理している情報まで入れてしまうと身動きが取れなくなってしまう。今回の調査でしか取り得ない情報や属性群が一番重要であり、経済産業省自体がクレンジングしたものの属性と、外部で実施したものをマージせずそのままグルーピング・紐付けすると、共通キーの提供や運用のコントロールがしやすくなる。
- ・法人番号では個人事業主が把握できないことについては、インボイス番号は本格的に運用されていくと、個人事業主もカバーできていくはず。中小企業やスタートアップを考えると法人番号の有用性追求、個人事業主はインボイス番号の把握が今後重要になってくるかもしれない。
- ・調査名簿は日本の事業所・企業の所在地や基本情報が入った情報の塊であり、調査を実施するために最も重要なものである。目的外利用であっても、名簿は本来積極的に貸し

出すことは避けたいものである。現状共通キーがないことにより、複数統計調査を繋げる必要があったり、統計内で時系列方向に繋ぐ必要があるという一定の要件が揃うと貸し出している状況。本事業で共通キーを作成することにより統計調査間の同一事業所・企業を接続することができるので、名簿提供を限りなく減らし、リスクを無くすることができる。提供の方法としては、各統計調査の個票に共通キーを付けるか、各調査の個票の ID とキーが入った別表を作成し提供するのがよい。つまり、経産省は名簿に共通キーを付与しても、利用者への提供時には名簿を渡すことなく、個票に附帯または、事業所 ID と共通キーの別表（コンバータ）の提供が望ましい。

- ・利用目的に沿って提供項目を絞って申請することについて、省内は交渉出来ると思うが、外部機関に対しては難しいと思う。なぜなら利用者は、後で必要になったときに再度申請するのが面倒だから最初から全部申請しておきたいと言ってくると思われるからで、結局全部出さなければいけなくなるのではと危惧する。
- ・最も慎重にならないといけないのが緯度経度情報で、この情報だけで事業所名称もその他情報も特定できてしまう。この緯度経度情報は名簿と同じくらいの機密性だと考えるべき。だからこそ、名簿が欲しいと言った人にしか渡さない、というような考えを基にした運用ではなく、そもそももう名簿は貸し出さない、共通キーは個票か別表にする、かつ共通キーの位置情報の貸し出しは厳密に必要な場合や、緊急性があると判断されたときだけ、などのルールを決める必要がある。
- ・緯度経度情報による位置情報については提供の条件を事前にしっかり決めることが大事。緊急かつ重要な事態でない限り提供しないという形で厳しめにすることが重要。位置情報はほぼ名簿情報と同じなので貸し出しには十分な精査が必要。審査が大変になるとすれば、現状各室で行っている提供業務について、全体をみて判断する部署や専属部署を設置することや、人材を確保することができればいい。
- ・100%正確ではないことを了解した人に使ってもらうことについては、共通キーを作って調査統計グループがサービスするのではなく、各者と協力してお互いにフィードバックしながら改善していくスタンスが重要。
- ・議論しているのは申請ルールだが、承認した後、データを繋いだ後の利用の仕方や禁止事項についてのルールも示していく必要がある。

以上