

第3回 国税庁保有行政記録情報の整備に関する有識者検討会

議事要旨

日時：令和5年5月29日（月）11:00～12:30

場所：Web開催

事務局から、配付資料に基づき、主に「国税庁保有行政記録情報の整備に関する技術検証ワーキンググループ」での検討結果と、データ提供の整備方針について説明。その後、以下のとおり御意見をいただいた。

- ・貸出方式は、研究者にとっても利便性の高い提供形態である。その一方、閲覧方式の可能性についても、引き続き検討をしていく必要がある。
- ・ビジネスデータに対する匿名加工は容易ではなく、海外でも非常に事例が少ない。まずは、事務局案とおり、パーソナルデータから開始して、次のステップでビジネスデータの議論をしてはどうか。
- ・利用者の研究目的によってニーズは異なるため、例えば特異値が想定されない研究においては、最初から特異値が含まれるレコードを除いた範囲からサンプリングを行うなどの方法も考えられる。
- ・全数ではなく標本抽出することはやむを得ない。標本抽出の検討に際しては、抽出割合よりも標本サイズの方が重要である。
- ・匿名加工については、研究者側のニーズも踏まえながら、加工の程度を考える必要がある。
- ・年齢と性別と地域の情報が全てであると、対象が絞られる可能性が高くなるため、いずれかは提供しない又は情報の粒度を粗くする必要がある。
- ・学術研究の目的のために提供するのであれば、その学術研究の範囲について検討する必要がある。その検討に当たっては「何が学術研究に含まれるか。」よりも「誰が行う学術研究を対象とするか。」がより重要な論点ではないかと考えられる。

- ・ 利用者の範囲については、学術研究の目的のために提供することを踏まえると、少なくとも大学や国立研究開発法人等の研究機関に属している研究者は対象とされるべきである。
- ・ データを利用した場合、そのデータを利用してどのような研究が実施されたかも公表するべきではないか。
- ・ 将来的には、オーダーメイド集計の1つの方法として、提供を受けたデータで分析を行った後、国税庁に研究で行った分析用のコードを送付し、国税庁において、データの作成元となるマザーデータで当該分析を実行してもらい、統計表や回帰結果を提供してもらう方法も考えられる。そうしたことも見据えて、提供するデータとマザーデータは同じ形式としておくべきではないか。
- ・ 利用規約等により利用者に課せられる守秘義務をどのように取り扱うかによって、提供可能なデータの粒度は異なってくる。その意味で、守秘義務に関する法的な観点からの整理が求められる。
- ・ サンプルデータは、テスト用なのかあるいは教育用なのか、さらには提供するデータに近いものと捉えるのか等のサンプルデータの位置づけの整理も必要である。
- ・ 総じて、今後の方向性は事務局が示した案でよい。今後も引き続き、検討・整備を進めていただきたい。

以上