

研究要旨

2023年度の分担研究では、全国がん登録情報の利用提供マニュアルにおいて講じるべき技術的な安全管理措置についての現状と課題の検討に資することを旨として、海外における公的統計に対する匿名化措置の現状や、二次利用の実態に関する調査研究を行った。具体的には、公的統計マイクロデータ、行政記録情報、及び医療健康データの二次利用の実態に関して、海外の統計作成部局やデータアーカイブ・データ提供サービス機関を対象に調査研究を行った。

本調査研究によって、海外の統計作成部局や医療健康データのアーカイブ機関で展開されているリモートアクセスの現状とリモートアクセスに基づく個票データの提供状況、さらには個票データに基づく分析結果の安全性に関するチェック方法といった個票データにおける匿名化措置の最新状況に関する有益な知見が得られた。本研究の成果として得られた個票データの利活用と匿名化措置に関する海外動向を踏まえつつ、全国がん登録情報の利用提供マニュアルにおいて講じるべき技術的な安全管理措置の現状と課題について、今後さらなる検討を行うことが求められよう。

A. 研究目的

欧米諸国においては、学術研究目的を指向する形で、公的統計のマイクロデータ及び行政記録情報を用いた実証研究が追究されてきた。特に、匿名化処理が施されていない公的統計や行政記録情報の個票データ（非識別データ（deidentified data））、さらには医療健康データにおいては、オンサイト施設のようなセキュアな環境の下で利活用が広範に進められてきた。近年では、新型コロナウイルス感染症の影響もあり、欧米諸国の多くの国でリモートアクセスの整備が検討され、研究者が、研究室等において、公的統計や行政記録情報の個票データ、さらにはリンクされた個票データに、オンサイト施設が稼働しない時間帯でもアクセスして実証研究を行うことが可能になっている。

このような海外の動向については、全国がん登録情報の利用提供マニュアルにおいて講じるべき技術的な安全管理措置を検討する上での有益な参考情報になり得ると考える。

本研究の目的は、海外の公的統計や医療健康データの二次利用の現状や展開の方向性について調査研究を行うことにある。それによって、全国がん登録情報の利用提供マニュアルにおいて講じるべき技術的な安全管理措置についての現状と課題の検討に資する有益な情報を得ることを指向している。

B. 研究方法

本研究では、全国がん登録情報の利用提供マニュアルにおいて講じるべき技術的な安全管理措置についての現状と課題の検討に資するための調査研究として、2023年度は、統計作成部局である欧州連合統

計局（Eurostat）、オランダ統計局（オランダ）とデンマーク統計局（デンマーク）、研究機関である National Health Observatory（ルクセンブルク）、医療健康データのアーカイブ・提供サービス機関である eSante（ルクセンブルク）、DHD（オランダ）、Danish Health Data Authority（デンマーク）を対象にした実地調査を行った。具体的には、公的統計マイクロデータ、行政記録情報、及び医療健康データの二次利用に関して、統計作成部局、研究機関、医療健康データのアーカイブ機関の担当者や専門家にインタビュー調査を行った。インタビュー調査においては、近年諸外国で、学術研究目的のための公的統計や行政記録情報の利活用の推進のために展開されているリモートアクセス（大学の研究室のような研究を行う場所で使用する PC から個票データが保管されているサーバにリモートでアクセスできる仕組み）、及びリモートアクセスを含むセキュアな分析環境で個票データを用いて実証分析を行った結果の秘匿性に関するチェック（output checking）の方法を中心に、最新動向についての調査を行った。

なお、本研究では、イギリス人口センサスの個票レコードに関する管理の状況やセンサスの個票情報の利活用の歴史的展開を追究するために、大英図書館（British Library）や The National Archives を訪問し、文献資料の蒐集だけでなく、関連資料に関する情報収集も行った。

（倫理面への配慮）

本研究で行ったインタビュー調査の内容をもとに研究成果をとりまとめを行うにあたっては、調査対象者に研究成果への使用に対する許諾を取っている。

C. 研究結果

本研究で実施した、統計作成部局、研究機関、医療健康データのアーカイブ機関の担当者や専門家にインタビュー調査を通じて、公的統計、行政記録情報、さらには医療健康データの二次利用に関して、学術研究目的に対する公的統計の個票データ、行政記録情報と医療健康データの利用サービスの現状及び個票データの秘密保護のための技術的な安全管理措置に関する最新の専門的知見を得ることができた。

特に、リモートアクセスの最新状況については、欧米諸国における多くの統計作成部局が、公的統計や行政記録情報の個票データを対象に、大学・研究機関の研究室からのリモートアクセスによる利用サービスを積極的に展開していることがわかった。またセキュアな環境における個票データの分析結果の外部への持ち出しに関するチェックの方法については、統計作成部局やデータ提供サービス機関における独自のガイドライン・基準にしたがって、担当職員がマニュアルで分析結果のチェックを行っているが、最近では分析結果に関する秘匿性のチェックの自動化についての運用の試みが図られている。

以下では、具体的な調査研究の成果として、Eurostat とデンマーク統計局を例に、公的統計や行政記録情報の提供に関する最新の動向について紹介する。

1. Eurostat

Eurostat においては、2002 年に学術研究用ファイル(Scientific Use File=SUF)の提供が開始されているが、事業所・企業系の統計調査では、Structure of Earnings Survey と Community Innovation Survey の 2 調査が、個票データとしての Secure Use File として、Eurostat のオンサイト施設(Safe Centre)で提供されている。なお、Eurostat のマイクロデータに関しては、認可された機関 (recognized organization)においてのみ個票データの提供が可能になっている。そのため、認可された機関の研究者及びその共同研究者であれば、その所属する大学や研究機関を通じて個票データにアクセスすることができる。

ルクセンブルク内の Eurostat のオンサイト施設に研究者が直接訪問し、Secure Use File にアクセスすることについては、場所的・時間的制約が存在するため、そうした制約に対する対応が Eurostat 側に求められていた。そこで、2022 年より、Eurostat において Secure Use File のリモートアクセスの運用が開始された。現在では、デンマークのコペンハーゲン大学を含む 2 か所のリモートアクセス施設でもアクセスすることが可能になっている。

Eurostat は、リモートアクセスによる個票データの提供にあたって、利用者自身が個票データを用いて分析した成果物の安全性に関して責任を負うことを重視している。そのため、Eurostat の担当者による分析結果の外部への持ち出しの際には、利用者側

がその安全性を事前にチェックした上で分析結果を持ち出すことの重要性を意味している。その一方で、提供者側も分析結果の持ち出しのチェックを行っているため、Eurostat においては、提供者と利用者が、分析結果の安全性について責任を分担していることが考えられる。

なお、西イングランド大学ブリストル校(University of the West of England Bristol)の研究グループによって、利用者自身がサーバからの持ち出しの対象となる分析結果に関する安全性の可否を自動的にチェックするためのソフトウェアである ACRO(=Automatic Checking of Research Outputs) が、2021 年に開発された。ヨーロッパ諸国の統計作成部局においては、分析結果のチェックのためのツールとしてまだ実用化されていないが、Eurostat では、試行的に ACRO を用いた分析結果のチェックを行うサービスを行っているとのことである。

2. デンマーク統計局

デンマーク統計局では、公的統計の作成の元になる行政記録情報を対象に、リモートアクセスを通じてのみ個票データの利用サービスが行われている。利用者は、人口社会情報、労働・所得情報、教育に関する情報、企業経済情報、医療健康情報といった様々な情報を対象に、各種のレジスター(行政上の登録簿)から、分析に必要な変数群を対象にリンケージがなされた個票データをリモートでアクセスすることが可能になっている。Eurostat と同様に、デンマーク統計局においても認可された機関に所属する研究者のみに対して、個票データのリモートアクセスが容認されている。

デンマーク統計局から提供された個票データで分析された結果については、デンマーク統計局が管理・運営しているスキャンシステムを通じて、研究者自身がサーバから外部に持ち出そうとする分析結果について自動的チェックを行うことが可能になっている。その一方で、一部の分析結果については、デンマーク統計局の職員がランダムに選別、事後的にマニュアルでチェックを行っている。

なお、近年のデンマーク統計局における個票データの利用サービスの動きとしては、マーケティングリサーチのような学術研究以外の研究目的に対しても、利用者に対して個票データの提供を行っている点が興味深い。ただし、このようなケースであっても、商用目的のために個票データを利用することはデンマーク統計局によって容認されておらず、対象となる研究が、社会に還元した研究、公共の利益に資する研究であることが求められている。

D. 考察

本調査研究で得られた海外動向をもとに、わが国における個票データのリモートアクセスの適用可能性を具体的に検討する場合に、個票データの申請者

が安全な利用者であり、利用者を含む研究プロジェクトが安全なプロジェクトであるかどうかを審査し、リモートアクセスによるデータの提供の可否を判断するプロセスが求められる。そのための法制度的面での具体的な議論が今後必要になるように思われる。その意味では、Eurostatやデンマークで展開されている、認可された機関に所属する者に対して個票データのアクセスを容認する仕組みについては、わが国の個票データの利用資格について議論する上で、参考になる点が少なくないように思われる。

リモートアクセスの場合、オンサイト施設のように利用場所を限定し、カメラ等で施設内の作業状況を逐次モニタリングすることは困難だと言える。そのため、個票データのリモートアクセスでの利用サービスにおいては、データの提供者側が安全なネットワーク環境を技術面で保証することが必須要件であって、そのための基盤整備は、リモートアクセスの運用に向けての要件だと言える。

統計作成部局やデータ提供サービス機関にとっては、研究者がオンサイト施設やリモートアクセスといったセキュアな環境において分析した成果物の外部への持ち出し審査にあたって、マニュアルで分析結果のチェックを行っている。そのために、分析結果をマニュアルでチェックすることができる職員の確保が求められる。こうしたことから、分析結果の自動的なチェックを可能にすることは、データ提供側にとっても重要な関心事項となっている。その意味で、ヨーロッパ諸国で開発されたACROによる分析結果の自動チェックについては、様々な分析結果に対するその適用可能性等、統計作成部局による実用化に向けた取り組みに関する動向を今後も注視する必要があると言える。

E. 結論

本研究では、海外の統計作成部局や医療健康データのアーカイブ機関で展開されている個票データリモートアクセスの現状と個票データに基づく分析結果の安全性に関するチェック方法を中心に、調査研究を行った、本研究の成果については、全国がん登録情報の利用提供マニュアルにおいて講じるべき技術的な安全管理措置についての現状と課題について検討を行う上で有益であると考えます。

本研究成果で得られた海外の事例も参考にした上で、わが国におけるリモートアクセスの運用可能性や分析結果のチェック方法・基準について今後検討することが求められよう。

F. 健康危険情報

(分担研究報告書には記入せずに、総括研究報告書にまとめて記入)

G. 研究発表

1. 論文発表

1) Ito, S., Terada, M., Kato, S. “The Potential of Differential Privacy Applied to Detailed Statistical Tables Created Using Microdata from the Japanese Population Census”, Paper presented at UNECE Expert Meeting on Statistical Data Confidentiality 2023, Wiesbaden, Germany, September 26th, 2023, pp.1-10.

2) 伊藤伸介・横溝秀始「わが国の公的統計における合成データの展開可能性に関する一考察—事業所・企業系の統計調査を例に—」『経済学論纂(中央大学)』第64巻3・4合併号, 147~164頁, 2024年1月10日

3) 伊藤伸介・寺田雅之・加藤駿典「公的統計に対する差分プライバシーの適用と有効性の評価に関する検討—国勢調査を例に—」『統計研究彙報』第81号, 69~88頁, 2024年3月29日

4) 伊藤伸介・木村映善・佐々木香織「イギリスにおける大規模データの二次利用の新たな展開—スコットランドの事例を中心に—」『統計研究彙報』第81号, 89~108頁, 2024年3月29日

2. 国内発表

1) 伊藤伸介「海外における公的統計と行政記録情報の利活用と秘匿措置に関する新たな動向」2023年度統計関連学会連合大会, 京都大学, 2023年9月5日

2) 伊藤伸介「公的統計における匿名化の有効性評価と合成データの生成技法に関する海外動向」, コンピュータセキュリティシンポジウム2023, アクロス福岡, 2023年10月31日

3) 伊藤伸介「リモートアクセスによるマイクロデータの提供に関する海外動向」公的統計マイクロデータ研究コンソーシアムシンポジウム2023, オンライン開催, 2023年11月24日

4) 伊藤伸介「海外における公的統計マイクロデータと行政記録情報の提供状況と匿名化措置」, 研究集会「大規模データの公開におけるプライバシー保護の理論と応用」, 統計数理研究所, 2023年12月7日

H. 知的財産権の出願・登録状況

無し