公的統計ミクロデータ研究コンソーシアム

NEWSLETTER

Vol.01

MAR.2022

INDEX

創刊の挨拶 …… 02

特集：創刊インタビュー …… 03

活動報告 …… 07

今後の活動予定 …… 09

問合せ先 …… 09

創刊の挨拶

藤井　良一

公的統計ミクロデータ研究コンソーシアム評議会　議長
(大学共同利用機関法人　情報・システム研究機構　機構長)

公的統計ミクロデータ研究コンソーシアム評議会の議長を務めます情報・システム研究機構の藤井良一です。 このたびの NewsLetter の創刊にあたり、 一言ご挨拶を申し上げます。

公的統計は、 国民にとって合理的な意思決定を行うための基盤となる重要な情報です。 そして、 公的統計作成の基礎となるミクロデータを分析することにより、 我が国の実情を多様な観点から客観的に評価する EBPM（エビデンスに基づく政策立案）の実現が期待されています。 しかしながら、 我が国においてミクロデータは価値の高いデータであるにもかかわらず、 その価値を十分に引き出せていませんでした。

そのような状況を踏まえ、 本コンソーシアムは公的統計ミクロデータの利用推進を目的に2016年に設立され、 産官学のメンバーが緊密に連携して研究利用の環境整備の推進に取り組んでまいりました。 特にミクロデータの利用拠点である「オンサイト施設」 を全国展開する活動に注力し、 当初 4 拠点だったオンサイト施設は現在14拠点に拡大するなど、 着実な成果が得られています。

また利用機会の少なさに伴う研究スキルの不足などにより、 ミクロデータの利用が少数の研究者に留まっている状況を改善するため、 ミクロデータ利用の普及活動に広く取り組んでまいりました。 統計行政の最新動向、 ミクロデータの利用事例を紹介するシンポジウムを毎年開催し、 今年度はコンソーシアム・ホームページをリニューアルし、 コンソーシアム概要の紹介記事、 ミクロデータ利用の具体的ノウハウに関するチュートリアル動画の掲載等、 コンテンツの大幅な拡充を行いました。 このような啓発活動の結果、 本コンソーシアムの会員数も着実に増加しています。 現在もコンソーシアムの評議会、 運営委員会では、 公的統計の利用に関する課題、 施策を活発に議論しており、 ミクロデータの利用制度改善に関する提言、 ミクロデータ利用者に対する実践的な内容のサービス提供を計画しています。

このたび、 今後さらに活動領域を広げる本コンソーシアムの情報を定期的にお伝えするため、 NewsLetter を発行することと致しました。 コンソーシアムの活動の成果、 得られた知見などを取り上げ、 公的ミクロデータをより広く活用していただくために有意義な情報発信をしてまいりたいと思います。 今後とも本コンソーシアムへの皆様のご支援・ご協力を賜りますことをお願い申し上げます。

座談会

公的ミクロデータ二次利用の道のりと活用促進への期待

公的統計ミクロデータ研究コンソーシアムの南和宏運営委員長と、コンソーシアム設立に深く関わった統計数理研究所の椿広計所長、山下智志副所長、岡本基主任 URA の 4 人が、当初のいきさつを振り返るとともに、将来の展望を語り合いました。

30年越しの官学ネットワークがコンソーシアムに結実

――みなさんが公的統計と関わるようになったのはいつ頃からですか？

南　和宏　統計数理研究所　教授

南　「私は2016年 5 月から統計センターの非常勤研究員となり、公的ミクロデータのオンサイト利用における安全性審査をするための基準づくりを担当しました。それが、このプロジェクトに関わることになったきっかけです。オンサイト利用制度がスタートした今も、週に一度、統計センターへ出向き、基準運用の助言をしています。

コンソーシアムの運営委員長は、2 年前に初代委員長の山下先生から引き継ぎました。」

椿　広計　統計数理研究所　所長

椿　「私の場合は1990年に、統計審議会調査技術開発部会の専門委員を命じられたのが最初でした。統計審議会は、統計行政の整備・改善について提言を行う総務庁長官の諮問機関で、現在の総務省統計委員会の前身です。

私は経済統計のことはよく知らなかったので、そこに来ていた統計局の人たちや若い統計学者のために経済統計の専門家である美添泰人先生（青山学院大学経済学部名誉教授）が主催していた勉強会に参加しました。公的統計の分野の人たちと付き合うようになったのは、そこからですね。」

南　「なるほど、その頃からの交流がコンソーシアムで実を結んだのですね。」

岡本　基　情報・システム研究機構
主任 URA ／特任准教授

岡本　「私は2010年に、情報・システム研究機構（ROIS）の新領域融合研究プロジェクトの技術補佐員として椿所長のもとで、当時統数研に所属していた元山斉先生（青山学院大学経済学部教授）と共に公的統計のオンサイト利用制度のスタート準備に携わったのが最初です。

それから10年以上ずっと公的ミクロデータに関わっていまして、2016年にコンソーシアムが立ち上がってからは、事務局の実働を担ってきました。」

山下智志　統計数理研究所　副所長／教授

山下　「私が公的ミクロデータのプロジェクトに参加した理由の一つは、それまでの自分の研究のスキームが活かせると思ったからです。

私の専門はファイナンス関係の統計学で、全国160万社の企業の財務諸表データを集めて解析に使えるようにクレンジングしたデータベースの構築・社会実装を手掛けました。それ以外にも、国際協力銀行（JABIC）や国際協力機構（JICA）といった政府機関のデータを使った研究の経験もあったことから、公的ミクロデータについても財務諸表データのように共有できれば面白い、と思っていました。

ただ、実際にプロジェクトに参加することになったのは、ROIS の中期計画によって2016年にデータサイエンス共同利用基盤施設が新設され、そこに統数研が中心となって社会データ構造化センターを立ち上げることが決まってからです。」

――新しくできるセンターで、以前から温めていた企画を実現しようと思われたわけですね？

山下　「最初に『公的ミクロデータをデータベース化し、利活用を実現することをセンターの目玉としよう』というアイデアを出されたのは椿所長でした。私はそれをスキームに落とし込む役割を担いました。

2015年の夏から半年ほどかけてROISや文部科学省など各方面と折衝を続け、北川源四郎機構長（当時）を議長とするコンソーシアムを設立しました。データサイエンス共同利用基盤施設の正式な事業と位置づけたことで、総務省の協力も得られましたし、統数研の存在意義をアピールする結果にもなったと思います。」

統計法の改正で二次利用への道が一気に拓けた

――コンソーシアムの設立以前、公的ミクロデータの利用はどんな状況だったのでしょうか。

椿　「欧米では公的ミクロデータが経済学の研究に利活用できるのに、日本ではまったくできないというのが実情でした。日本の経済学の先生方は、1995年頃からこのことに問題意識を持っていました。

1996年に、一橋大学経済研究所の故・松田芳郎先生が代表者、美添先生が分担者となり、大型の科研費研究が立ち上がりました。その中で公的ミクロデータの利活用のあり方が５年間、検討されたのです。

当時の統計法では公的ミクロデータの二次利用はものすごく大変だったので、一橋大学が包括的に『目的外使用申請』を行い、いわゆる『プログラム送付型のオンサイト』の試行をしました。研究者がプログラムを送ると、一橋大学で解析を行い、結果を返してくれるものです。

また、匿名データや疑似データの作成方法などもいろいろと研究しました。私も美添先生の手伝いで、貯蓄動向調査と家計調査のリンケージをしたり、合成データを作ったりしました。当時は利用促進について研究者の間では大いに盛り上がっていたんです。研究成果は『統計情報活用のフロンティアの拡大の総括的研究』というタイトルで出版されました。」

――この研究の成果が出たことで、公的ミクロデータの利用は進んだのですか？

椿　「いや、だからといってすぐに制度が変わるわけではありません。統計法の壁は厳然とあり、私も自分で作った合成データを『これはデータから作った乱数であって、本物ではありません』と主張したけれども認めてもらえず、お蔵入りになってしまいました。総務省の政策統括官室に『これを公開できるかどうかの審査には20年かかるだろう』と言われましてね（笑）。

オンサイト利用に向けて再び動き出したのは、2013年ぐらいからですね。2009年に統計法の改正によって『統計データは国民全体の資産である』と位置づけられたことが大きかったと思います。これによって『研究のための二次的な利用』が再浮上しました。

当時東京大学教授だった廣松毅先生が座長を務めた『統計データの二次的利用促進に関する研究会』など

さまざまな研究会ができ、二次利用に向けて大きく動き始めました。

当時の川崎茂応用統計学会長が日本学術会議にオンサイト拠点の構想をマスタープランとして提出。これが認められたことを根拠として、2014年に統計審議会の後継である内閣府統計委員会が、公的統計の整備計画を閣議決定まで持ち込んだのです。」

椿　広計　統計数理研究所　所長

――長い道のりでしたね。

椿　「いや、まだ続くのです（笑）。コンソーシアムができた2016年の時点でも、まだオンサイト利用の法整備は完了していませんでした。オンサイトできちんと利用できるようになったのは、改正統計法が施行された2019年５月からです。

総務省はその前年、和歌山県に『統計データ利活用推進センター』を設置し、オンサイト利用に向けて整備を進めていました。そうしてさまざまな角度から検討し、ようやく法改正が実現したのです。

私も一度、総務審議官に呼ばれて『公的ミクロデータがオンサイトで使えるようになったら、ノーベル経済学賞ぐらい取れるようになるのか』と聞かれたことがあります。これには『少なくとも、ミクロデータが使えなければノーベル賞など取れないことは確かでしょう』と答えました。」

データサイエンスへの熱視線を二次利用促進のエンジンに

――コンソーシアムができて６年たちましたが、公的ミクロデータをめぐる状況は、どのように変化したでしょうか。

岡本　「『二次利用できる』ということが研究者の間に少しずつ浸透し、当たり前のことになりつつあると感じます。オンサイト拠点を設置した大学はまだ10校ほどですが、2020年度からはシンフォニカ（公益財団法人統計情報研究開発センター）が、オンサイト環境整備に必要な経費の助成事業を開始しましたから、費用面は心配ありません。

データサイエンスに力を入れる大学は確実に増えており、私が講義を受け持つ大学では経済学部の学生の８割がデータ分析を履修しています。すでに種まきはできた、という状況でしょう。」

山下　「コンソーシアムを取り巻く社会の環境を含めて考えると、この６年間でデータサイエンスへの注目度は飛躍的に高まり、好循環が生まれていると感じます。

データが有効利用されるためには、データそのものの整備、それを使いこなすことのできる人材、そしてそれらを支える予算という三つが揃ってはじめてうまく回るわけです。

われわれはこれまで人を集め、データを使える形にして、細々とではありますが、情報が流れる仕組みを整えてきました。今はそこにきちんとした予算や政策、大学のアクションなどが加わり、大きな流れになっています。コンソーシアムは、この潮流の中の一つの重要な要素と位置づけられています。追い風が吹き続けている印象です。」

――現状の課題と展望は。

南　「コンソーシアムの会員数は徐々に増えていますが、研究者はまだ十分といえるほど多くありません。われわれのメッセージが届いていない学術領域が、まだ残されているのではないかと見ています。これまでは経済学が中心でしたが、それ以外にもミクロデータが役立つ領域はたくさんあるはずなので、そうした方面への広報活動に力を入れていくつもりです。」

椿　「そのためには、公的ミクロデータを使ってこんな解析ができるというユースケースをどんどん出していく必要がありますね。

統計データ利活用推進センターが置かれた和歌山では、県がミクロデータを利用して政策決定に役立てているという例もあります。データを使いこなせる人材さえいれば、自治体にとっても有用でしょう。」

南　「今はまだ、オンサイト拠点に敷居の高さを感じる研究者も少なくありません。そこで、まずは若手研究者や大学院生を対象に、疑似データを用いてハンズオンのトレーニングをするワークショップを開催しようと、計画しているところです。

私も大学で講義を持っていますが、公的統計は学生にも馴染みのあるものが多く興味が湧きやすいので、教

材にも取り入れています。そうやって学部の段階から公的統計に触れる人が増えてくれば、中長期的にはオンサイトの利用者も拡大するはずです。」

岡本　「今後は統数研としても、データサイエンス教育を担当する教員の育成に力を入れる方針です。そういう人材が育てば、それぞれの大学でデータサイエンスをさらに広めてくれると期待しています。」

椿　「符号表や欠測値の扱いなど、公的データには特有の慣習があり、それが利用者のハードルを上げていました。しかし、今、データの提供元である総務省でも、普通の人が当たり前に使えるエクセルや CSV の形に変更するなどの作業を進めています。」

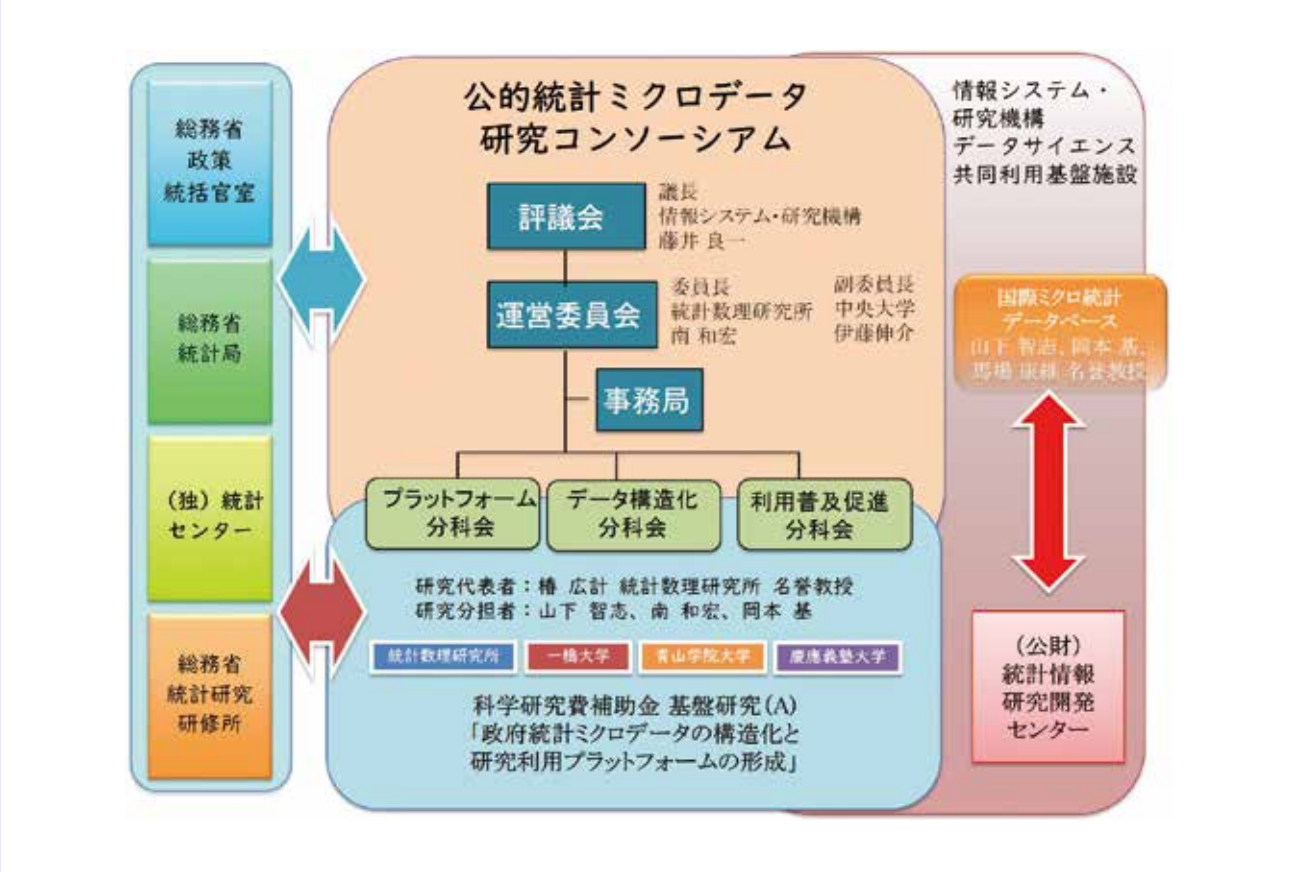
南　「調査を行う省庁や部署によって、例えば性別など共通の項目でも変数名が異なっていたりするのも標準化したほうがいいですね。単純なことでも、利便性に大きな影響があります。」

椿　「分散型のほうが政策にフィットしたデータが取れるメリットがありますから、これまではそうなっていたのです。価値の高い分析をするには、統一するほうが望ましいわけですが、難しい問題です。」

山下　「仕様の統一は、どの世界でもデータ作成・運用の一番の懸案ですね。私が関与した銀行のデータベースも、マニュアルが完全に統一できるまで20年かかりました。理想は共有されていても、実際の整備になるとしんどいことですからね。

しかし、データ活用に対する社会の理解がこれまでになく高まっている今、データ提供側と利用者側がそれぞれの課題を共に乗り越える機は熟したと言えるでしょう。」

――官学民の連携事業であるコンソーシアムが、その牽引車となってくれることを期待しています。本日はありがとうございました。



コンソーシアムでは、公的統計の利活用に関するステークホルダーと密に連携して活動を行っている。データ提供側の行政機関から公的統計に関する最新動向の情報を入手するとともに、データの利用者である会員の要望を行政機関へ伝える。

Report 1 公的統計ミクロデータ研究コンソーシアムシンポジウム2021開催

2021年11月19日（金）に「公的統計ミクロデータ研究コンソーシアムシンポジウム2021」を開催しました。本シンポジウムは、大学や研究所などの公的統計ミクロデータを学術研究に利用する関係機関とデータの提供側である総務省統計局や統計センターなどが連携して2016年から毎年開催しており、今回で６回目を迎えます。

昨年に引き続き、新型コロナウィルスの感染拡大をうけオンラインでの開催となりましたが、通常では参加が難しい遠方の地域からの申込みもあり約100名の方々にご参加いただきました。

本シンポジウムでは、午前のセッションは「オンサイト利用に関するチュートリアル」と題し、公的統計ミクロデータの提供側・利用側の双方からの講演をいただき、今後のオンサイト利活用の促進に繋がる実践的な内容になりました。

また午後の部では、冒頭に藤井良一（本コンソーシアム評議会議長、情報・システム研究機構長）より開会の挨拶があり、その後２つのセッションで講演が行われました。

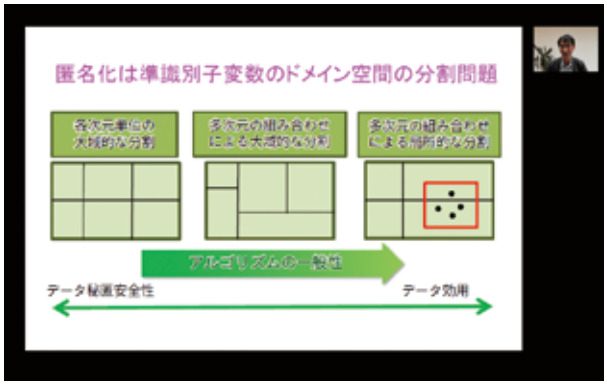
「公的統計と統計教育」のセッションでは、人材育成が急務といわれている統計教育について、その最新の取り組みと動向に関する４件の講演があり、また後半の「統計データの高度利用に関する研究」のセッションでは、疑似ミクロデータの作成やデータの匿名化に関する３件の講演が行われました。シンポジウム全体を通して活発な議論が行われ、有意義なシンポジウムとなりました。

http://jmodc.org/event/sche_repo_2021sympo.php

プログラム

〈オンサイト利用に関するチュートリアル〉

- 「統計データ利活用センターにおけるオンサイト利用推進の取組」
  赤谷 俊彦（総務省統計局・独立行政法人統計センター 統計データ利活用センター）
- 「オンサイト利用による分析結果等の安全性確認における注意と事例」
  阿部 穂日（独立行政法人統計センター 統計情報提供課）
- 「公的統計匿名データを利用したデータサイエンス講義のための取組み」
  白川 清美（立正大学 データサイエンス学部）



〈公的統計ミクロデータ研究コンソーシアムシンポジウム〉

第１セッション「公的統計と統計教育」

- 「政府における統計人材の確保・育成に向けた取組」
  稲垣 好展（総務省 統計局 調査企画課）
- 「データサイエンス教育と公的ミクロデータに期待する役割」
  渡辺 美智子（立正大学 データサイエンス学部）
- 「統計エキスパート人材育成プロジェクトの推進」
  千野 雅人（統計数理研究所）
- 「SSDSE を中心とした統計教育への取組み」
  山下 雅代（独立行政法人統計センター）

第２セッション「統計データの高度利用に関する研究」

- 「Synthetic Data の考え方に基づく疑似的なミクロデータ作成の可能性」
  高部 勲（立正大学 データサイエンス学部）
- 「公的統計データの匿名化に関する海外の動向とわが国における課題」
  伊藤 伸介（中央大学 経済学部）
- 「公的ミクロデータに対する k- 匿名化加工の検討」
  南 和宏（統計数理研究所）

Report 2 第 9 回評議会開催

定款に定められた手続きに従い2021年12月14日（火）に、第 9 回評議会をオンラインで開催しました。この評議会では下記の議案が提案され、いずれも承認されました。

第 1 号議案：評議員の推薦について
第 2 号議案：運営委員の交代について
第 3 号議案：第 6 事業年度活動報告
第 4 号議案：第 7 事業年度活動計画
第 5 号議案：コンソーシアム会員の認定について

第 1 号議案では、會田雅人氏（公益財団法人 統計情報研究開発センター専務理事）の評議員就任が承認され、第 2 号議案では、田中雅行氏（一橋大学経済研究所准教授）、中村英明氏（総務省政策統括官室（統計制度担当）付　参事官）、赤谷俊彦氏（総務省統計局統計データ利活用センターセンター長）の運営委員就任が承認されました。

総会当日の様子

その後、データサイエンスをテーマにした他の機関、コンソーシアムとの連携、教育分野での公的統計ミクロデータの活用促進、データ利活用に関するチュートリアルやワークショップの実施、広報活動の更なる展開などについて活発な意見交換が行われました。

Report 3 デジタルの日創設記念　特設ページ開設

人に優しいデジタル化の実現を目指す「デジタルの日」の趣旨に賛同し、コンソーシアムの活動を紹介する特設ページを公開しました。

http://jmodc.org/digitalday.html

本特設ページでは、公的統計ミクロデータ研究コンソーシアムの紹介として、設立の経緯やコンソーシアムはどのような組織か、現在の活動や今後の活動予定などコンソーシアムの概要ついて分かりやすく説明しています。

Report 4 統計関連学会連合大会企画セッション実施

2021年度統計関連学会連合大会において、「公的統計ミクロデータにおけるさらなる利活用をめぐって」と題する企画セッションを2021年 9 月 7 日（火）に実施しました。このセッションでは、公的統計における疑似データの活用、世帯類型と働き方の差が消費構造に与える影響、公的統計ミクロデータを使用した宿泊需要の季節変動分析、宿泊割合の推定と可視化など、 5 つの研究発表が行われました。来年度も公的ミクロデータの利用を啓蒙する企画セッションを検討しています。

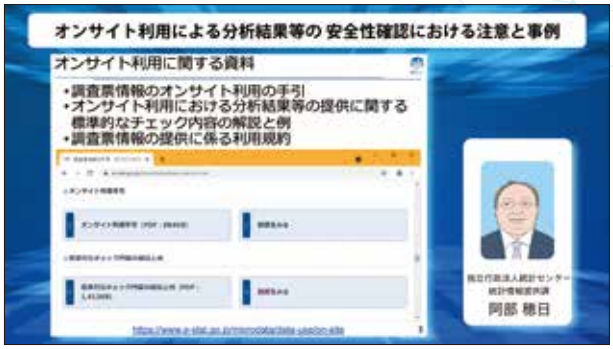
Report 5 シンポジウムのチュートリアル動画公開

2021年11月19日（金）に開催された「公的統計ミクロデータ研究コンソーシアムシンポジウム2021」のうち、下記２件の講演はオンサイト利用に関して実践的で有用な情報を提供しており、今後も広く参照できるようにチュートリアル動画としてウェブサイトで公開します。

オンサイト利用に関するチュートリアル Web サイト

http://jmodc.org/videos/

●「統計データ利活用センターにおけるオンサイト利用推進の取組」
赤谷 俊彦（総務省統計局・独立行政法人統計センター 統計データ利活用センター）



●「オンサイト利用による分析結果等の安全性確認における注意と事例」
阿部 穂日（独立行政法人統計センター 統計情報提供課）

活動予定

公的統計ミクロデータ利活用ワークショップ開催

来年度、公的ミクロデータの研究利用に興味がある方を対象に、ワークショップを開催します。第１回目は、国の調査票情報等の利用制度の概要、主な調査票情報の内容の紹介、疑似的なミクロデータを使ったデータ分析の実習を行う予定です。詳細が決まりましたら公的統計ミクロデータ研究コンソーシアムのウェブサイトにてお知らせいたします。

http://jmodc.org/

尚、今後開催するワークショップのテーマに関する希望がございましたら、事務局までご連絡ください。

統計関連学会連合大会企画セッション

来年度も統計関連学会連合大会にて公的統計ミクロデータ利活用に関する企画セッションを予定しております。詳細が決まりましたら、ウェブサイトにてお知らせいたします。

問合せ先

ADDRESS

〒190-8562　東京都立川市緑町10-3
大学共同利用機関法人　情報・システム研究機構
データサイエンス共同利用基盤施設　社会データ構造化センター内
公的統計ミクロデータ研究コンソーシアム　事務局

MAIL office@jmodc.org
URL http://jmodc.org/