

公的統計マイクロデータ研究コンソーシアム

NEWSLETTER

INDEX

Vol.06

OCT.2024

特集：東京大学
社会科学研究所インタビュー 02
調査票情報利用のための
オンサイト施設紹介 07

活動報告 08
今後の活動予定 09
問合せ先 09

座談会

社会現象の解明に欠かせない 社会調査データと公的統計マイクロデータ

出席者：
東京大学 社会科学研究所 附属社会調査・データアーカイブ研究センター
准教授 (社会調査研究分野) 准教授

藤原 翔さん、石田 賢示さん

東京大学 社会科学研究所 比較現代社会部門 准教授

永吉 希久子さん

聞き手：
中央大学 経済学部 教授、公的統計マイクロデータ研究コ
ンソーシアム 運営委員会 副委員長

伊藤 伸介さん

公的統計マイクロデータはさまざまな分野の研究に活用されている。経済学や社会学もその一つだ。今回は、東京大学社会科学研究所^(注)に所属するお三方に、社会学の分野で公的統計マイクロデータがどのように役立っているかを伺った。同研究所が公開する社会調査データ「SSJデータアーカイブ」との性質の違いや使い分けも必読。聞き手は経済統計学を専門とし、当コンソーシアムの運営委員会副委員長である中央大学の伊藤伸介教授が務めた。



伊藤 伸介 教授



藤原 翔 准教授



石田 賢示 准教授

同じ「社会学」でも、 アプローチは多種多彩

——東京大学の社会科学研究所（通称・社研）は、社会科学の諸分野を横断的に結びつけた共同研究を実践する研究所で、社会調査の個票データを蒐集・保存して公開する「SSJデータアーカイブ（SSJDA）」を運営しています。皆さんの分野は社会学ですが、どのような研究をなさっているのでしょうか。

藤原 「大学院のときから社会調査を専門とする研究室に所属し、チームで社会調査を実施したり、SSJDAに寄託されたデータを二次利用したりして、計量分析を行ってきました。テーマは『社会階層論』です。今は、親と子の職業の関連、またそこに教育や学歴が果たす役割などを分析することで、現代日本社会の格差や不平等の実態を明らかにする研究をしています。最近は、こうしたテーマに因果推論や機械学習をどう適用で

きるかにも関心を持っています。」

石田 「私は大学院で『教育社会学』を専攻し、学生の就職活動、すなわち社会学の言葉では『学校から職業への移行（トランジション）』とありますが、これを研究していました。ただ、就職してもその後に転職することもあるので、今は30代ぐらいまでを見据えて若者のキャリア形成や、転職でチャンスを得る場合とそうでない場合の違いに関する要因などを分析しています。また『移民のルーツ



永吉 希久子 准教授

を持つ子どもの教育問題』にも取り組んでおり、公的統計マイクロデータでは厚生労働省と文部科学省共管の『21世紀出生児縦断調査』を二次利用させていただきました。」

永吉 「私は藤原さんと同じ研究室の出身で、当時から『外国人に対する態度』の研究をしています。対象国による態度の程度の違いや、否定的な態度が向けられやすい要因などを日本の社会調査データや、ISSP (International Social Survey Programme) 国際比較調査のデータなどを用いて分析しています。最近SNSに投稿されたツイートなどのビッグデータを使ったり、実験と組み合わせた調査を行ったりと手法を拡大しています。また、外国籍の人たちの生活や職業、賃金の状況とその要因なども研究しており、公的統計では国勢調査などのデータを使っています。」

——「社会学」というバックグラウンドは同じでも、皆さんそれぞれにアプローチの仕方が異なることが分かり、興味深く伺いました。

社会現象の実態を明らかにし、メカニズムを追求する

——そもそも社会学というのは、どういった対象や課題について探求する学問なのでしょう。また、社会学においてなぜデータを使って明らかにすることが必要なのかも教えてください。

藤原 「社会学とは何か——。100人の社会学者がいれば100通りの答えがありそうですが、私なりに説明すると、基本的には社会現象の実態を明らかにし、そのメカニズムを追求する学問です。社会現象とは例えば、生まれ育った家庭によってその後の社会経済的資源の獲得、つまり所得、学歴、職業、結婚の機会や、ライフスタイルなどが異なるといったことです。社会調査や公的統計は、そうした実態を明らかにするのに適した方法だと考えています。」

——経済学の分野においては、経済現象を明らかにするのが目的の一つだと考えますが、経済学ではすでに確立された理論の中で議論がなされる傾向があります。社会学のアプローチは、そうではないように感じました。

石田 「社会学の中にもいくつかの立場があるとは思いますが、人間観というか、人間がどう考え、どう行動するかという想定は、経済学に比べて自由かもしれません。また、社会学では『集団』に注目する傾向があります。使うのは個人を対象としたデータでも、そこから見ようとしているものは、その個人によって構成されて

いる集団、あるいはその集団間の関係なのです。つまりデータを使って、集団や集団間の関係を取りまく社会問題を量的に可視化する。社会現象の実態とメカニズムの把握には、フィールドワークと計量的な分析を相互補完的に用いる手法が有効です。」

——伝統的な経済学では「合理的な個人」を想定します。合理的な個人が市場メカニズムによって最適な行動を取ることでモノ・ヒト・カネの動きが決まってくる、というのが経済学の考え方の基礎にあります。ただ、それだけでは実態を説明できない部分もあって、そこは個別具体的に考えなければいけない。そのとき着目するのが集団です。集団の現象をデータによって明らかにしようというモチベーションは経済学にもあり、社会学と重なる面がありますね。

永吉 「社会学の古典的な研究でも、国別の自殺率と宗教との関係を比較することで、社会の中に存在するパターンを明らかにするといったものがあります。社会学がパターンや規則性に着目する研究である以上、データを使うのは必然かもしれません。ただ、経済学と少し違うのは、社会学では意味世界に基づいて説明するところです。例えば、『人的資本』を考えると、何が人的資本に当たるかは所与ではなく、それ自体も人々の相互作用の中で決まってくる。つまり、理論的な前提を所与とせず、今、人々がどのように認識しているところを出発点としつつ、その背景に起こっていることを検証する。それが社会学の特徴ではないかと思っています。その意味では、データだけ

では最終的に説明しようとすることに到達できないこともあります。]

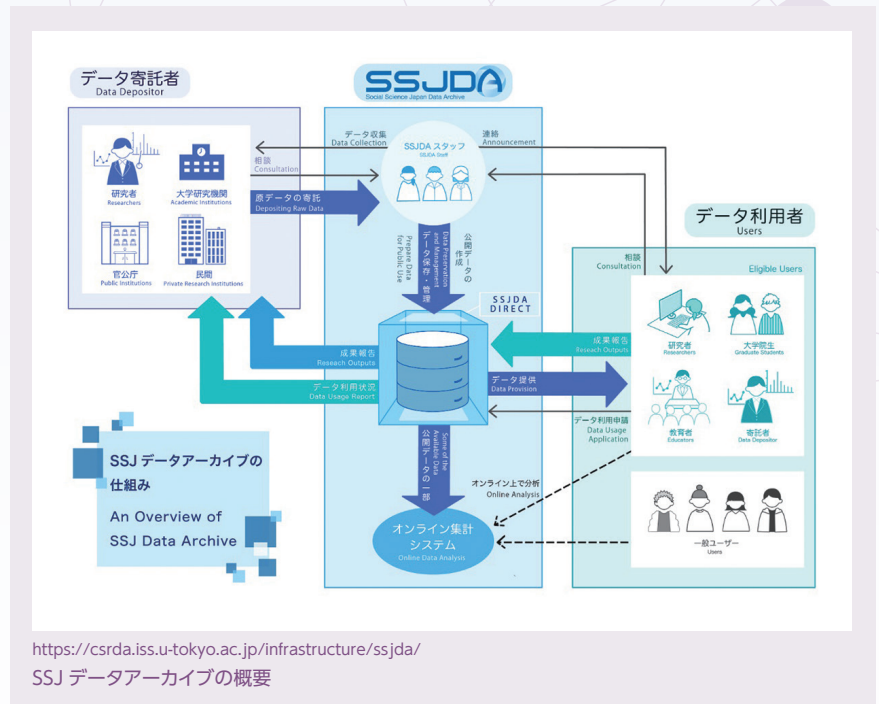
——経済学の場合、人的資本の評価に関しては、市場メカニズムを前提としたフレームワークの中で、賃金との関連で、議論がなされています。それに対して、社会学においては市場メカニズムでは捉えにくい分野にアプローチしているということですね。

一次分析と二次分析の両輪でオープンサイエンスに寄与する SSJDA

——皆さんが社会調査データを利用するとき、独自に調査票を設計して調査を行い、その結果を一次分析する場合と、SSJDAのデータを二次分析する場合があると思います。それぞれのメリット・デメリットは？

永吉 「二次分析の場合は、ほかの研究者が作った調査票が基になるので、すでにあるデータの中で何ができると考えます。ですから、まったく新しい問いを立てるような研究は難しいですね。自分で一から調査を設計する場合は、これまで訊かれていない質問を入れたり、これまでアタックされていない対象をターゲットにしたりできますので、自分の仮説を検証しやすくなります。」

石田 「私は、一次調査データを作ってSSJDAに寄託する業務を担当しています。一方で、自分の研究のために調査を設計して一次分析をすることもあれば、SSJDAのデータを使って二次分析をすることもあります。



自分で調査をする資金がなかった院生時代に、きちんと設計された調査データを使わせてもらったのは、とてもいい経験でした。今でも、独自調査を企画する前にアーカイブから公開されている調査の質問票を参考にしたり、分布を見たりします。

二次分析の場合は、どうしたら元の調査と差異化できるか工夫は必要ですが、調査にかかる手間暇がいらぬぶん、分析方法の検討などに集中しやすいですね。ほかの研究者の二次分析の論文を見ても、統計モデルの選択や推定の方法など、かなり高度な分析をしている人もいて、その部分でオリジナリティを出しているようです。」

藤原 「私自身は独自調査を主に行っていますが、調査票を設計するときには、ほかの研究者が二次利用をすることを前提とした項目も盛り込んでいます。項目が増え過ぎると調査対象者の負担になり、焦点がぼやけてしまう面もありますが、なるべく二次分析につながるような形で社会調

査を実施し、そのデータをアーカイブに寄託し、それがさらに次の社会調査に結び付くといったサイクルをイメージしています。」

——SSJDAが公開されて25年以上たちますが、その果たしてきた役割をどう見えていますか？

藤原 「今言ったようなサイクルができてきているのは、SSJDAがあったからこそで、社会科学分野のオープンサイエンス化に向けてなくてはならない存在になっていると思います。2023年度には、累積公開データセット数は約2700に及び、年間の利用申請は1500件を超えました。」

石田 「データの散逸を防ぐことも、研究コミュニティにおける大きな使命でした。例えば1970年代に実施された全国調査のデータが残っていれば、50年前と現在を比較することもできます。寄託データや研究成果の蓄積も増え、最近では国際誌に掲載さ

れた論文も寄せられており、世界へ向けた情報発信にも貢献しています。」

——社会科学研究所では毎年「二次分析研究会」を開催されています。どんな活動ですか？

藤原 「私たちのセンターでは、SSJDAの共同利用・共同研究促進のための事業として、『参加者公募型研究会』と『課題公募型研究会』の二つを実施しています。前者は、特定のテーマに関する調査に参加者を募集して行う調査です。参加者には、どんなデータを使って、どんな研究をしたいかを提案して応募していただきます。年度末には研究成果報告会を開催し、広くオーディエンスとも議論します。

また後者は、研究者自身が課題やテーマの設定から用いるデータまでを設定して社研の研究者とともに研究会を組織し、共同で二次分析を行います。多くの研究会は年次報告書を作成し、その後に論文投稿や一般書籍の執筆といったプロジェクトへ繋がっています。」

公的統計は頼れる「信頼性の高い大規模データ」

——ここからは、公的統計マイクロデータの利用について伺います。皆さんがマイクロデータを使うようになったきっかけや、これまでどんな研究に利用してきたかをお話ください。

藤原 「社会学では職業に関する評価基準として『職業威信スコア』が知られていますが、それ以外の尺度を

探していたときに、公的統計である『就業構造基本調査』のデータを使えば、海外で頻繁に用いられている尺度（社会経済指標、SEI）を作れるのではないかと思いつきました。ただ、利用の仕方がわかりませんでした。たまたまお会いした伊藤さんに相談し、先ほど言った課題公募型研究会を組織していただき、共同で公的統計マイクロデータの申請をしたのが最初です。」

——そうでしたね。私も覚えていません。石田さんはいかがですか。

石田 「私は2013年の参加者公募型研究会で、『就業構造基本調査』の匿名データを使わせていただきました。その後、藤原さんと同じ課題公募型研究会に参加したときは、『社会生活基本調査』の調査票情報を使いました。そういう経験の中で、調査票情報やレイアウト、符号表などの読み方も勉強させてもらいました。それからは、先ほどの『21世紀出生児縦断調査』や国勢調査も使っています。

永吉 「私が公的統計を利用できると知ったのは、皆さんが参加した課題公募型研究会の成果を何かで見たのがきっかけでした。その後、慶應大学教授の竹ノ下弘久さんのプロジェクトで国勢調査のデータを使わせていただき、それからは自分で申請して『賃金構造基本統計調査』などを使っています。」

——公的統計のマイクロデータを利用するメリット・デメリットはどういった点にあるとお考えですか？

永吉 「私は研究テーマが外国籍の

方の生活状況に関するものなので、十分な分析ができるほどのサンプルサイズを確保できているデータは社会調査にはありません。その点、公的統計はとても大きいデータで、しかも全国を対象にしているので、非常に大きなメリットがあると思います。とはいえ、外国籍者の生活状況を知るために実施している公的な調査は現状ではほとんどないので、自分が知りたいこと、例えば、日本語能力や在留資格、国籍といった重要な変数が、必ずしも含まれていない点がデメリットにはなっています。」

石田 「規模の大きさは最大の魅力ですね。さまざまなセグメントというか、条件や属性に分けて細かく分析ができますから。しかも、一定の信頼性が担保されている。デメリットはあまり感じませんが、強いて言えば、公的統計は所管省庁がターゲットとする政策課題に関する統計表を作るのが目的なので、人々の主観や意識などに関する項目が調査票に含まれていないことでしょうか。そういう場合は社会調査データを使えばいいのですが、両者の性質の違いには留意する必要があります。」

藤原 「先ほど言った職業の尺度を作るときにも、大規模なデータはありがたかったですね。職業ごとにかかなりのケースが確保でき、信頼性の高い尺度を構築できました。一方で、親の職業と子どもの職業の関係を分析するときには、対象者である子どもの現状は分かっても、どういう家庭に育ったかといった過去のことは分からない。それはデメリットかもしれません。ただ、公的統計の多くが世帯全員を調査対象としているので、同居している親子であれば、子と親

の状況を同時に分析できる。これは大きなメリットです。公的統計ならではの枠組みで、私自身、分析の幅が広がりました。』

省庁間の規格統一や利用期限の延長を希望

——最後に、公的統計マイクロデータを利用する上で制約を感じることや、統計作成部局に対する要望があればお話しください。

藤原 「事前に研究計画をしっかりと固めて申請しなければならない点は、少し困ることがあります。まあ、データの無い状態でどの変数が必要かを検討し、分析プランをじっくりと考える機会を与えられているのだと、ポジティブに捉えています(笑)。それから、早期化を検討されていると聞きますが、手続きにかかる時間が短くなると嬉しいです。」

石田 「匿名データを使うとき、出生年などがカテゴリーにまとめられていたりすると、研究テーマによっては使いづらい場合があります。秘匿性の観点からは仕方ない半面、サイズが小さくてもいいから自由に使えるといいなと思うこともあります。符号表やレイアウトは本当によくできていて、あれがあるおかげで自分の分析するデータセットをすぐに作れるのでありがたいです。ただ、省庁によって規格が統一されていないために、困惑することがあります。』

永吉 「現状の利用環境で十分満足ですが、あえて挙げるなら、利用の期限を緩やかにしていただけたらと思います。論文を投稿して修正を求められたとき、すでにデータにアク

セスできなくなっていることもあるので、数年の猶予を許してもらえると助かります。また、カテゴリー数の多い調査で、レイアウト表で細かい値の定義が省略されていて、探している情報が見つけれずに困ったことがあり、コードブックのようなものがあると嬉しいです。』

——コンソーシアム運営委員会としては、ユーザーのご意見を公的統計マイクロデータ二次利用の運営に反映していきたいと思います。今日は社会学とは何かというご説明から、SSJDAの特徴まで、貴重なお話をありがとうございました。

(注) 東京大学社会科学研究所は、東京大学に11ある附置研究所の一つで、設立は1946年。比較現代法、比較現代政治、比較現代経済、比較現代社会、国際日本社会の五部門に加え、社会連携研究部門、附属社会調査・データアーカイブ研究センター(CSRDA)で構成されている

<https://jwww.iss.u-tokyo.ac.jp/>

調査票情報利用のためのオンサイト施設紹介

「オンサイト施設」は、情報セキュリティが確保された環境で、許可を受けた研究者がマイクロデータ（調査票情報）を用いて独自の集計・分析を行うための専用室です。本 NewsLetter では、総務省統計局等が中心となり、全国の大学・研究機関等への設置展開を進めている調査票情報のリモートアクセス型利用のためのオンサイト施設をご紹介します。

〈第1回：情報・システム研究機構〉

情報・システム研究機構では、立川地区に2019年に竣工した「データサイエンス棟」にオンサイト施設を設置しています。弊施設は、利用室と控室を分離しており、スペースにゆとりを持たせた施設構造となっております。利用室内には通信機器等の情報機器を持ち込むことが出来ないため、控室には利用者向けPCや作業デスク等を設置し、利用室での分析と並行して、施設内で情報機器を利用した作業等出来るように配慮しております。



情報・システム研究機構 データサイエンス棟

利用室には、分析ブースを2式整備しており、同時時間帯に2チームの利用予約を受け入れることを可能としております。設置機器は、デスクトップベースで構築しており、片方のブースに27インチデュアルモニタ、もう片方のブースに34インチウルトラワイドモニタをそれぞれ設置し、利用者それぞれの好みや作業内容に合わせ、環境を選択できるようにしているのが大きな特徴です。

立川地区の敷地内には弊機構の宿泊施設「赤池ゲストハウス」があり、ゲストハウスに逗留しながら、短期集中で分析に取り組まれる方もいらっしゃいます。また、標準的な開室日時は平日10-17時となっておりますが、短期集中分析等の利用をサポートするために、時間外・休日利用についても、柔軟に対応させていただいており、実際にご利用いただいている方からも好評をいただいております。

弊施設の詳細、ご利用・ご予約に関するお問い合わせ等は、弊施設ウェブサイトをご参照ください。



情報・システム研究機構のオンサイト施設

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 オンサイト施設

住 所 東京都立川市緑町10-3 情報・システム研究機構 データサイエンス棟 2階

開室日時 平日10時～17時 ※土曜日、日曜日、国民の祝日、12月29日～1月3日、及び、予約状況カレンダーにて休業告知されている休業日を除く ※休日・時間外利用については応相談

ウェブサイト <https://csds.rois.ac.jp/onsite/>

統計関連学会連合大会企画セッション

本年度の統計関連学会連合大会が、9月4日（日）～7日（木）に東京理科大学で開催されました。当コンソーシアムでは、「公的統計の二次利用における最新動向」をテ

マにした企画セッションを行いました。各報告では活発な議論が展開され、公的統計の二次利用に対する参加者の高い関心がうかがえました。

プログラム

オーガナイザー 南 和宏（統計数理研究所）、伊藤 伸介（中央大学）、高部 勲（立正大学）
座 長 山下 智志（統計数理研究所）

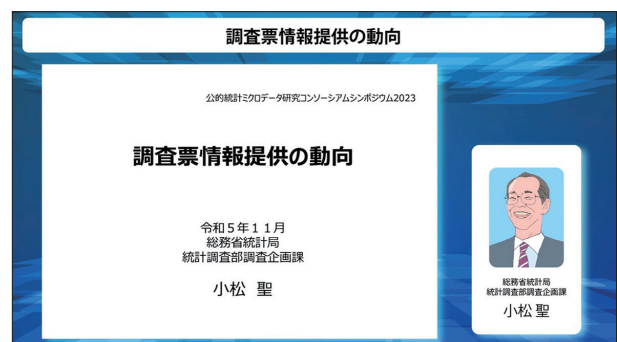
1. 「調査票情報提供の動向」
小松 聖（総務省）
2. 「統計的マッチングの公的統計へのさらなる検討」
高部 勲（立正大学）
3. 「わが国の公的統計における匿名データの作成可能性に関する実証研究—国勢調査のマイクロデータを用いて—」
伊藤 伸介（中央大学）、高部 勲（立正大学）、
村田 磨理子（統計情報研究開発センター）
4. 「海外における公的大規模データのリンケージの動向」
伊藤 伸介（中央大学）
5. 「公的統計における大規模統計表のプライバシー保護の最新動向」
南 和宏（統計数理研究所）



統計関連学会企画セッション 当日の様子

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動画公開

2023年11月に開催された「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」の講演動画をコンソーシアム公式ウェブサイト（<https://jmodc.org/videos>）で公開しました。今回は、総務省統計局 調査企画課の小松聖課長による「調査票情報提供の動向」を公開しております。本講演は、調査票情報の利活用を行う上で、大変興味深い内容となっております。ぜひご覧ください。



「調査票情報提供の動向」
小松聖（総務省統計局 統計調査部調査企画課長）

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2024

コンソーシアムでは、公的統計マイクロデータを学術研究に利用する大学や研究所の他、総務省統計局や統計センターなどの関係機関と連携し、シンポジウムを年1回開催しています。今年度も2024年11月19日(火)に「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2024」をオンラインで開催いたします。

プログラムの詳細やお申し込み方法については、近日中に公的統計マイクロデータ研究コンソーシアムの公式ウェブサイト (<https://jmodc.org>) にてご案内いたします。

連携イベント_公的統計マイクロデータ利活用に関する研究集会

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の連携イベントとして、前日の11月18日(月)に独立行政法人 統計センター主催の「公的統計マイクロデータ利活用に関する研究集会」がオンラインで開催されます。

この研究集会は、公的統計マイクロデータ等の利用者に限らず、公的統計を用いた実証研究・統計教育全般を対象としており、毎年様々な分野の方々からの研究報告や講演が行われております。同研究集会の詳細については、(独)統計センターのウェブサイト (<https://www.nstac.go.jp/use/archives/event/>) をご確認ください。

社会学分野でのマイクロデータ利用推進に関するワークショップ

社会学分野でのマイクロデータ利用を推進するために、公的統計および社会学の両分野の研究者が情報交換、議論を行うことを目的とした「社会学分野でのマイクロデータ利用推進に関するワークショップ」を2025年2月に対面形式で開催いたします。

前半は、データ利活用の制度やデータ基盤整備に関するチュートリアルセッションを実施し、後半では、公的統計や社会学の調査データを用いた研究利用の事例報告を行います。また、これまでのマイクロデータ利用に関するこれまでの経緯、現状を総括し、今後の課題と方向性を議論するパネル討論を予定しています。ぜひ皆様のご参加をお待ちしております。

日時：2025年2月14日(金) 10:00~17:00

場所：東京大学山上会館 大会議室(予定)

主催：公的統計マイクロデータ研究コンソーシアム

共催：東京大学社会科学研究所 附属社会調査・データアーカイブ研究センター

※プログラムの詳細は、決まり次第 コンソーシアム公式ウェブサイトでご案内いたします

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
公的統計マイクロデータ研究コンソーシアム 事務局

MAIL office@jmodc.org

URL https://jmodc.org/