

公的統計マイクロデータ研究コンソーシアム

NEWSLETTER

INDEX

Vol.04

OCT.2023

特集：公的統計マイクロデータと統計教育Ⅱ ～大学における統計教育～	02	活動報告	07
イベントレポート	06	今後の活動予定	08
		問合せ先	08

座談会

データドリブンで新たな価値を 生み出せる人材の育成へ

立正大学データサイエンス学部 教授

大井 達雄さん、白川 清美さん、高部 勲さん

司会：統計数理研究所データ科学研究系 教授

南 和宏さん

2022年に創立150周年を迎えた立正大学。2021年4月には、熊谷キャンパス（埼玉県熊谷市）に全国の大学で4番目となるデータサイエンス学部を設置した。統計センターで公的マイクロデータの公開に携わった白川清美さんと高部勲さん、研究・教育に公的マイクロデータを活用してきた大井達雄さんに、同学部における統計教育のいまとこれからの聞いた。

実業と地域に密着した 実践的な学び



南 和宏さん

南 「まず、立正大学データサイエンス学部の特徴をご紹介します。」

高部 「当学部は、文系・理系どちらにも開かれたカリキュラムとなっていることが最大の特徴です。『ビジネス』『観光』『社会』『スポーツ』の4つの分野を柱として、データサイエンスを実践的に活用する方法を学べます。1・2年次には一般教養と、専門分野に分かれる前段階のそれぞれ



高部 勲さん

の基礎を叩き込みます。その間に学生は自分がどの分野に適しているかを判断して3年次からゼミに入り、卒業研究論文を作成することになります。私と白川先生は公的統計、大井先生は観光や不動産と、いずれも文系寄りの科目を担当していますが、文系のデータサイエンスを教える大学はまだほとんどありません。

また、地域密着型の教育を掲げ、地元企業へのインターンシップや熊谷市と連携し、地元でのフィールドワークなどにも力を入れています。さ

らに、実務家教員が多いことも、当学部の大きな特徴です。プロパーの大学教員と実務家教員がだいたい半々の構成というのは珍しいと思います。」

南 「白川さんと高部さんは、これまで実務で公的マイクロデータに深く関わって来られましたね。」

白川 「私は大学院修了後に総理府統計局、現在の統計センターに入り、小売物価統計調査や全国消費実態調査などの調査から入力、結果表の作成、またその品質の審査までを担当



白川 清美さん

しました。その後、統計技術を研究する部署へ異動し、公的マイクロデータの二次利用を手掛けるようになりました。

立正大学では、データベースを専門に教えています。またこの秋からは、匿名データを使った分析や、SASプログラムを使った演習もスタートします。』

高部 「私も白川先生と同様、国家公務員でしたが、2年前に退職して立正大学に着任しました。実務で統計調査を実施していた経験から、講義では統計調査法や社会調査の設計・実施、また機械学習を使った分析で論文を書いたこともあるので、AI入門などを担当しています。」

白川 「私のゼミでは、起業するのに必要なことは何かを含めて考えています。例えば、スポーツサイエンスの分野で、アスリートの体にセンサーを付けてデータを取り、さまざまな分析をするといった事業について、スポーツ担当の教員と一緒に研究しているところです。」

また以前、JICAの仕事でカンボジアへ行っていたことがあり、現地の統計局とコネクションがあります。そこで今、立正大学の国際交流センターと、来年3月に学生をカンボジアにインターンシップに行かせる計画を立てています。どんな産業があり、どんな調査が行われているか、現地での起業も視野に、自分が学んできたことを何に生かせるかなどを考えてほしいと思っています。」

南 「大井さんは、ずっと大学人として研究・教育に携わって来られたのですね。」

大井 「はい。大学・大学院と経営学を専攻していたのですが、大学院の



ときの先生が土地統計を研究されていたので、私は建物も含めて不動産統計を研究することにしました。具体的には、企業不動産や住宅費、住宅取得能力などです。

その後、最初に赴任した大学が国際学部観光学科だったことから、観光学に興味を持ち、観光統計も研究するようになり、立正大学でも教えています。ですから、白川先生、高部先生と違って、ずっと公的マイクロデータを使う立場で社会科学、社会経済統計の研究をしてきたということですね。」

南 「先ほど、熊谷市の連携の話が出ましたが、大井さんのゼミでも何か活動をしていますか?」

大井 「そうですね。今年はフィールドワークの一環として、『熊谷うちわ祭』の来場者にアンケートを実施しました。一人あたりお金をいくら使っているかなどのデータから祭の経済規模を分析するなど、毎年継続すれば市にとっても有益な提案ができるのではないかと考えています。」

また、熊谷市はスマートシティーに取り組んでおり、この面でも今後、教員や学生のデータサイエンスの知見を生かした連携ができると考えて

います。」

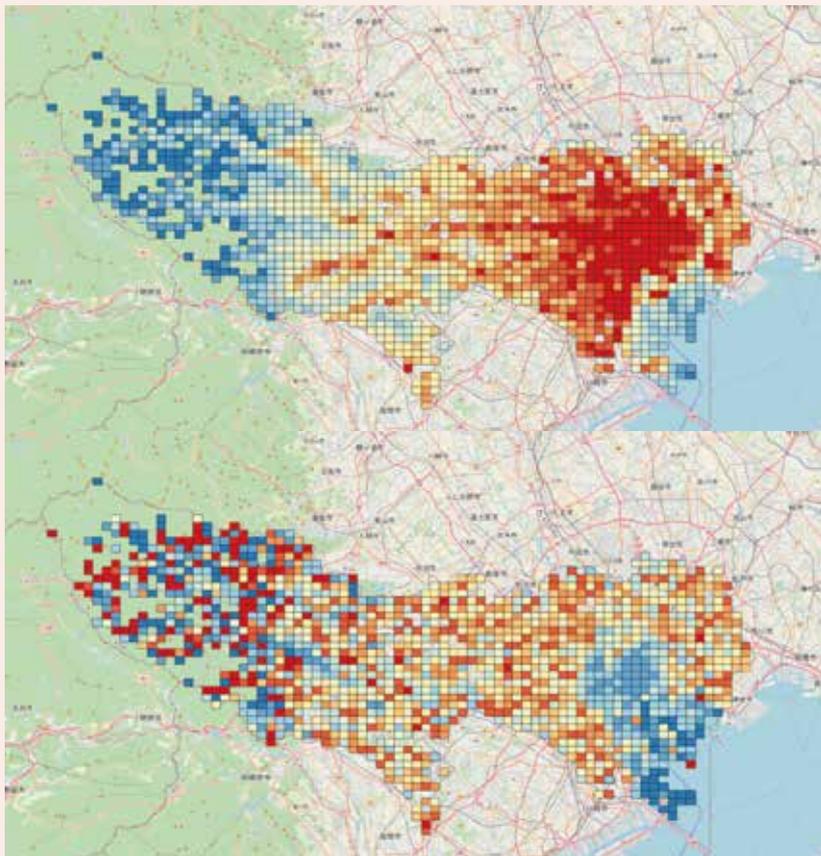
ビジュアル化やメタバース利用で学生に興味を持たせる

南 「皆さんそれぞれのご経歴を生かしてデータサイエンスを教えていらっしゃるわけですが、授業では学生に興味を持ってもらうために何か工夫をされていますか? 白川さんは匿名データをどんなふうに使われるのでしょうか。」

白川 「匿名データはオンラインで使ってはいけないので、オフラインの演習室でSASを使って教えます。また、GUIで分析できるSAS Viyaを使ってAIモデルを構築するなど、学生が喜びそうな分析手法も教えていきたいと思っています。こちらはオンラインでプログラムを動かすので、教育用データセット『SSDSE』を使う予定です。」

もう一つは、メタバースの中で、アバターを使ってデータを取ることをやってみようと思っています。メタバースでゲームをやっている若者は多いので、そういう身近な世界で調査や分析に慣れてもらえれば、と。」

大井 「私の講義では、G空間情報センターで無料公開されている47都道府県の月別人流データや、観光庁が公表している観光客のデータを入手し、GISソフトで分析しています。例えば、2019年の東京都の滞在人口の分布では、東側の都心部に滞在人口が集中していますが、2020年のデータを見ると、都心部は減り、西側の郊外で増えています。その理由を『コロナ禍によって在宅勤務が増えたから』と推測するといった内容です(図)。」



東京都における1kmメッシュ別の滞在人口データ(上が2019年、下は2020年)
出典：「全国の人流オープンデータ」(国土交通省) (<https://www.geospatial.jp/ckan/dataset/mlit-1km-fromto>)

ただ、いきなり分析といっても学生は何を見ているかわからないので、マーケティングや経済学、観光統計などの基礎知識を並行して教えながら、シナジー効果で理解できるように工夫しています。]

南 「観光というジャンル自体が、文系学生にも身近ですね。」

大井 「学生には『観光産業はこれからますます伸びる。イノベーションは成長産業から起こるのだから、今こういう勉強をしておく、社会に出てからビジネスチャンスをつかめる』という話をよくしています。それに、観光産業は裾野が広いので、観光についてのデータや分析が飲食産業や流通産業などにも応用できる、ということも少なくありません。」

文系学生はイメージができないと理解できない傾向があるので、データ分析の結果を数字ではなく、ビジュアライズして直感的に把握できるようにするなど、苦心しています。]

南 「高部さんは、社会調査の講義ではどんなことを教えていますか？」

高部 「調査票の作り方を含めて、企画から実施までをひと通り教えます。実務での経験から、まず『調査票はいったん配ったら後戻りできないから、最初にしっかり考えて作らないといけない』ということを2年次に教え、3・4年次では実際にアンケートを作って配付し、回収したデータをクリーニングして分析する、という流れです。」

また、立正大学にも公的マイクロデー

ータを取得できるオンサイト施設ができたので、そこを使った研究例なども学生に紹介しています。『公表された統計結果だけでなく、その裏にあるデータを使うと、こんな分析もできる』と伝えるのがねらいです。例えば、労働力調査の結果は失業率や就業者数などとして公表されていますが、元になった個票データには住所や収入など1世帯ごとの情報が詰まっている。これらを使えば、他にもさまざまな分析ができる、ということを教えています。]

知名度向上と申請の簡素化が公的マイクロデータ普及の鍵

南 「公的マイクロデータを『提供する側』と『使う側』の両方を経験して、何か気づいたことはありますか？」

高部 「自分で使ってみて感じたのは、『なんて申請手続きが面倒なんだろう』ということ(笑)。でも、最近はずっと改善が進み、申請書も作りやすくなったし、審査期間も短くなりましたね。それでも初めて使う人は戸惑うかもしれないので、コンソーシアムに入ってもらいたいかもしれません。それがコンソーシアムを立ち上げた目的の一つでもあります。私自身、提供側にいた頃からコンソーシアムで意見を言ったり、聞いたりしながら使用ルールなどの改善に役立てていました。]

南 「白川さんは、公的マイクロデータの利用マニュアルとも言える『使ってみよう！マイクロデータ』の改定にも携わられたそうですね。」

白川 「はい。せっかく調査表情報の二次利用ができるようになったのに、

申請の仕方やマイクロデータに付随する『符号表』の使い方がわからないことがネックになって、利用が進まないのはもったいない。そこで、自分でテキストを書こうと思ったら、すでにシンフォニカ（公益財団法人 統計情報研究開発センター）から発行されていたんです。そこで、同センターと協働して2018年の改正統計法に即した内容に改定し、マイクロデータの活用事例などもアレンジして、2023年3月に出版しました。同センターの機関誌『エストレーラ』にも、『公的統計匿名データの利用とその分析』と題して隔月連載で寄稿していますので、コンソーシアムの皆さんにも参考にしてもらえればと思います。」

南 「公的マイクロデータを大学教育に活用するうえで、どんな点に課題を感じていますか？」

大井 「授業で使う匿名データの下準備をすべてわれわれがやらなければいけないのは辛いですね。分析よりも下準備のほうがずっと時間がかか

るわけですし、その過程で学ぶことも多いので、学生自身が準備をできるように、データを扱う権限を多少緩めていただけると助かります。」

高部 「公的マイクロデータの秘匿性を高めた匿名データは、『情報量は少ないが使いやすい』という触れ込みでしたが、実際は利用のハードルはマイクロデータとほとんど変わりませんね。もう少し秘匿性を強めてもかまわないから、使い勝手がよくなるように制度を改めるほうがいいと、私も思います。」

白川 「審査を行う統計センターの現場も、申請者とのやりとり非常に手間がかかってパンク寸前で、許可までの時間が長びく一因ともなっているんです。もう少し基準を緩めないと、申請する側も許可する側も大変ですね。」

高部 「私が感じているもう一つの課題は、公的マイクロデータを使った研究例の少なさです。デジタル社会の基礎知識を身につける『リテラシーレベル』の教育では、『実データ』を使った分析内容を紹介することが重視されています。公的マイクロデータはまさに実データですから、一番いいのは、教員が実際に公的マイクロデータを使い、その成果を教材として学生に教えること。ところが、研究者の中にも公的マイクロデータの存在さえ知らない人がかなりいるのが実情です。本物のデータに触れる喜びを多くの人に知ってもらうには、まずは公的マイクロデータの認知度を上げることが先決かもしれません。」

実データの分析スキルを武器に各界で羽ばたく

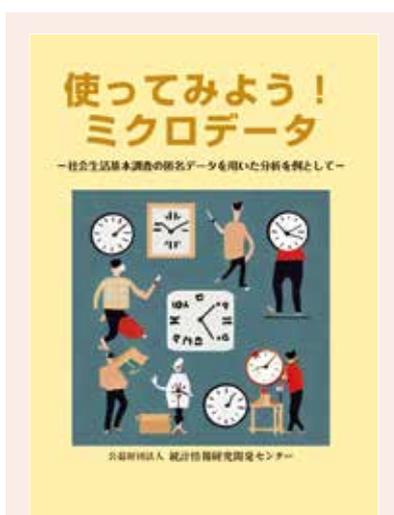
南 「立正大学データサイエンス学部では、2025年3月に最初の卒業生を輩出します。日本の次世代を担う若者たちにどんな期待をしますか？」

大井 「これからは、理系・文系関係なく、あらゆる人がデータを使いこなして、世の中に新しい価値を生み出していく時代です。日本にはまだまだデータ人材が足りません。当学部の卒業生たちがリーダーとなって、日本にイノベーションを起こしてくれることを願っています。」

高部 「データを扱える人材があらゆる業種に浸透していき、その職場で評価され頼られる存在になってほしい。実社会で本物のデータに直面したとき、当学部で匿名データや公的マイクロデータの研究例に触れた経験を生かして活躍してくれることが、究極の目標です。」

白川 「データドリブンで思考できる人材は、転職を繰り返してステップアップしていくこともできるだろうし、起業という選択肢もあるでしょう。われわれも、そのためのスキルをしっかりと教え、若者たちが自分のやりたい職業に就けるようにしてあげたいと思っています。」

南 「貴重な実データである公的マイクロデータを大学教育に有効活用することで、日本のデータサイエンスのボトムアップを実現させていきたいものですね。本日はありがとうございました。」



「使ってみよう！マイクロデータ—社会生活基本調査の匿名データを用いた分析を例として—
(公益財団法人統計情報研究開発センター刊)
<https://www.sinfonica.or.jp/publish/irregular/other.html>

2023データ利活用シンポジウム in 高野山 - 地方創生 - データサイエンスの薦め

2023年8月26日（土）、和歌山県の高野山大師教会中講堂にて、総務省、独立行政法人統計センター、和歌山県主催による「2023データ利活用シンポジウム in 高野山 - 地方創生 - データサイエンスのすすめ」が開催されました。

本シンポジウムは、地域DXの実現に資する観点から、行政や社会に流通する各種データを活用した政策立案の先進事例を通じて、データの利活用の推進を和歌山から全国へ向けて発信する主旨のもと、今回が6回目の開催となります。

シンポジウムは、総務大臣、和歌山県知事、来賓ご挨拶ののち、講演が始まりました。

はじめに、マッキンゼー・アンド・カンパニーの工藤卓哉氏より、基調講演「生成AIがもたらす経済効果と将来展望」が行われ、生成AIの行政での利用について、氏の経験や産業界での活用事例を紹介しつつ、生成AIによるバーチャルアシスタント、レポート生成、原因特定などを挙げ、それらの活用可能性について言及されました。

次に、東京都市大学の秋山祐樹氏より、特別講演「自治体保有データと政府統計を活用した空き家分布推定の取組」が行われ、氏が自治体と取り組んでいる政府統計データ等を活用した空き家調査の効率化に関する研究事例を紹介し、自治体・企業が持つデータ・統計の新たな使い道・使い方の可能性、自治体におけるDX推進のためのマインドトランスファーの必要性について述べられました。

休憩を挟み、パネルディスカッションが開催され、コーディネータの椿広計氏（統計数理研究所）、パネリストの工藤氏、山内慶太氏（慶應義塾大学）、浜口知実氏（日本生命）が登壇されました。それぞれの立場や研究成果の紹介を含めたポジショントークののち、地方創生におけるデータのあり方などについて、意見が交わされました。最後に椿氏が、AIの成長だけでなく、人間自体がデータとともに成長すること、課題解決・対話のためのデータ利活用の重要性などを挙げ、シンポジウムは盛況のうちに終了しました。



公的統計マイクロデータのためのチュートリアル・講習会 動画公開

2023年1月24日に開催した「公的統計マイクロデータのためのチュートリアル・講習会（講義形式）」の動画を公開しています。これらの動画は、公的統計マイクロデータに馴染みのない方、まだ実際の利用に至っていない潜在的なユーザーにむけた、公的統計マイクロデータの概要についての入門的な内容となっております。

今回は総務省統計局 赤谷俊彦氏による「公的統計の概要」、(独)統計センター 千葉亮太氏による「二次的利用の制度」、総務省統計局 小田大輔氏による「2020年基準改定における世帯消費動向指数（CTI ミクロ）の合成方法の改善」の3本の動画を公開しました。

チュートリアル動画ウェブサイト

<http://jmodc.org/videos>



本コンソーシアムではこれまで、公的統計の制度や概要、オンサイト施設の利用、施設開設に伴う手続きなど、計7本のチュートリアル動画を公開しています。また今後も公的統計マイクロデータについて理解を深めていただけるよう、順次動画を公開予定です。

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3日(日)～7日(木)に京都大学にて開催され、本コンソーシアムでは「公的統計のデータ構造化とマイクロデータ分析の展開」と題し

た企画セッションを実施しました。当日は46名の参加者があり、各報告とも更なる公的統計マイクロデータの利活用につながる活発な議論が展開されました。

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)、高部 勲 (立正大)
座長 山下 智志 (統計数理研究所)

1. 「公的統計マイクロデータにおけるデータリンケージに関する課題と展望」

村田 磨理子¹、伊藤 伸介²、出島 敬久³
(1. 統計情報研究開発センター 2. 中央大 3. 上智大)

2. 「データリンケージ・マッチングの手法を用いた公的統計マイクロデータの更なる活用可能性について」

高部 勲 (立正大)

3. 「公的マイクロデータの正規化とデータベース設計の検討」

南 和宏 (統計数理研究所)

4. 「日本生まれの子どもたちの教育達成：21世紀出生児縦断調査による親の国籍のあいだでの比較」

石田 賢示 (東京大)

5. 「国勢調査データを活用した国際結婚における同類婚パターンの分析」

永吉 希久子 (東京大)



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

コンソーシアムでは、公的マイクロデータの利活用を推進するため様々な活動に取り組んでいます。2016年以降、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。シンポジウムでは、オンサイト利用を中心に公的統計の行政制度の最新動向や公的マイクロデータを利用した研究事例などを紹介しています。

本年度も2023年11月24日（金）10時～16時40分（予定）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2023」をオンラインで開催いたします。

午前の部では「公的統計オープンデータに関するチュートリアル」と題し、総務省統計局／統計センターによるチュートリアルその他、一般用マイクロデータやSSDSE（教育用標準データセット）についての講演を予定しています。これから公的統計マイクロデータの利用を検討している方だけでなく、既に利用している方などにも向けた、マイクロデータの利活用を支援するチュートリアルとなる予定です。

また、午後の部では「公的統計における二次的利用の新展開」「統計をめぐる諸課題に関する研究」と題して、2つのセッションを実施予定です。

本シンポジウムは事前申し込み制（無料）のオンライン形式のシンポジウムとなっております。プログラムの詳細、申込み先のURLは10月初旬に公的統計マイクロデータ研究コンソーシアムのウェブサイト (<https://jmodc.org/>) にてお知らせいたします。ぜひ、皆様のご参加をお待ちしております。

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/>