

社会と統計

第6号

社会情報教育研究センター開設10周年記念号

立教大学 社会情報教育研究センター 研究紀要

2020年2月

社会と統計

第6号

社会情報教育研究センター開設 10 周年記念号

【巻頭言】

- ・ 社会情報教育研究センター開設 10 周年に寄せて 松本 康 (3)

【社会情報教育研究センター10 周年記念座談会】

- ・ CSI のこれまでとこれから
 - － 開設メンバーで振り返る 10 年間のあゆみ－菊地 進 松本 康 山口 和範 水上 徹男 (5)

【論文】

- ・ ミクロデータ分析と公的統計データベースに関する展開
 - － 2019 年度出張報告を兼ねて－櫻本 健 西林 勝吾 濱本 真一 (25)
- ・ データサイエンス教育の現状と課題
 - － 高等教育機関におけるデータサイエンス教育の方向性－山口 和範 山口 誠一 門田 実 (45)

【研究ノート】

- ・ 『日本帝国統計年鑑』を電子化しての雑感 五十嵐 彰 (55)
- ・ 質的データアーカイブ構想の現状と課題
 - － 数値化されていない調査データの保存と活用に向けて－高橋 かおり (65)

【資料】

- ・ 海外データアーカイブの動向 3
－ IASSIST 年次大会の報告から －

五十嵐 彰 高橋 かおり (75)

【センター活動報告】

- ・ 2019 年度社会情報教育研究センター研究活動等報告 (83)

《巻頭言》

社会情報教育研究センター開設 10 周年に寄せて

立教大学社会情報教育研究センター (Center for Statistics and Information: CSI) は、2010 年 3 月に、「調査・情報・統計技法の活用による研究活動の高度化への寄与」および「学部学生・大学院学生に対する研究基礎能力の涵養」を目的とする学内の教育研究支援組織として設置され、2020 年 3 月に開設 10 周年の節目を迎えました。この間、社会調査部会、政府統計部会、統計教育部会の 3 部会体制で、社会調査士・専門社会調査士および統計検定・統計調査士等の資格取得支援、社会調査法、政府統計の二次利用および統計分析法に関する教育支援コンテンツの開発や学部科目・全学共通カリキュラムへの協力、コンサルティング業務、外部講師を招いてのセミナーや研究会の開催、社会調査データアーカイブ (RUDA) の運営など、多彩な事業を展開してきました。さらに、国内外との対外連携事業にも積極的に取り組み、内外にその存在感を高めてきています。

近年、データサイエンス教育の必要性が広く認知されるようになる一方で、一部の政府統計の信頼性が社会問題化するなど、調査・統計への社会的関心が高まりつつあります。そのため、大学における調査・統計に関する基礎教育の重要性もあらためて認識されてきているところです。また、欧米ではすでに確立しているデータインフラとしてのデータアーカイブの構築・整備も、ようやく日本でその必要性が認識されはじめてきています。

もともと CSI 設立の出発点となったのは、欧米では確立している社会調査・統計関係の教育研究支援組織が、日本ではほとんど見られないという認識からでした。この 10 年間に国内の他大学でも少しずつ、このような組織の必要性が認識され、設立される例が見受けられるようになりました。その背景には、情報技術革命の進展にともない、大量の情報を収集・分析できるようになり、調査技法・統計技法を適切に評価し、運用する能力がますます求められるようになったことが挙げられます。

こうしたなかで、本号では、CSI の運営に深くかかわってきた教員が集まり、CSI の 10 年をふりかえり、その意義をあらためて確認し、今後を展望し課題を語り合う座談会を掲載しました。これをひとつの節目として、今後も、CSI の意義が広く理解され、その機能が十分に発揮されることを願ってやみません。

松本 康

《社会情報教育研究センター10周年記念座談会》

CSIのこれまでとこれから —開設メンバーで振り返る10年間のあゆみ—

菊地 進・松本 康・山口 和範・水上 徹男

I 2009年のセンター開設準備

水上: 今日はお集まりいただきまして、ありがとうございます。社会情報教育研究センター、Center for Statistics and Informationの座談会を開催させていただきます。本センターは調査、情報、統計技法を活用した教育研究支援の組織として、2010年3月に開設されました。現在10周年を迎えたところでございます。センターには、3つの部会があります。政府統計の利活用の仕方について支援する政府統計部会、次に社会調査の実施を支援する社会調査部会、3つ目に統計リテラシーの習得と向上を支援する統計教育部会です。部会活動だけでなく、社会調査士ですとか、統計検定など、社会調査に関係した資格取得の支援活動を広く展開してきました。学部への授業提供も行ってきています。例えば、調査関係の授業として、27コマあります（同一科目が含まれるため、コマ数で提示）。これらは全学共通科目の中の基礎科目と大学が10学部あるうち8学部提供している科目が2科目。大学院への科目提供がありまして、2020年度の調査関係授業の提供が27コマとなります。さらに調査分析のコンサルティングも行っています。

どのような経緯で本センターができたかなど、本日、これまでに多大な貢献をしていただきました先生方に、開設時の状況から現在の問題、さらに将来構想についてお話いただきます。経済学部菊地進先生、社会学部松本康先生、経営学部山口和範先生にお越しいただきましたので、まず開設の準備につきまして、菊地先生からお願いできますでしょうか。



菊地 進

菊地: 「社会調査センター設置要望」、これがやはりスタートでしたので、後で、松本先生からそのところを詳しくお聞きしたいと思っています。2009年度の4月半ばのことです。総長室に「社会調査センター設置要望」が出されているというのを聞きました。私はそのころ総長補佐をやっていたので、当時の総長室長から取りまとめの担当をやって欲しいとの依頼を受けました。それで、5月に社会調査センター設立検討チームというのを立ち上げて、ここに松本先生、山口先生、それからコミュニティ福祉学部の坂田周一先生にも加わっていただきました。

このチームで検討を始め、数次の会合を経て、6月18日には部長会に「社会調査センター設立検討チーム答申の方向性（協議）」を出しています。この時点で新しい組織につい

てかなり踏み込んだ話をしています。実は、この部長会協議に先立つ5月末ごろ、国から大型補助金の募集があるという話が伝わってきておりまして、それをにらみながらこの答申の方向性を出したという経緯があります。ただ、部長会協議の時点では、まだ大型補助金の申請に社会調査検討チームの提案内容が学内的に採択されるか、まだはっきりしていませんでした。ですから、いくつか候補がある中で学内採択に残していく取組みと、それから検討チームの答申を詰めていく取組みを同時並行で進めていました。

そうしたところ、6月末には、学内でこのセンター構想を上回る内容で応募できる場所はないということがわかってきました。そこで思い切った申請案を提出し、学内採択を得ることができたわけでありまして、松本先生が書かれていますように、これでセンター設置構想と補助金の構想とが実現に向けて一気に動き出したわけでありまして。

夏休みに入る前の7月末の部長会で実は助教の人事枠の承認を得ています。ですから、そのころには文部科学省から採択通知が届いていたはずですが、8月10日に設置準備室会議を開いていますが、ここですでに新任人事を具体的に進める議論がなされています。補助金に関わる支援体制整備事業の計画という資料も残っていますが、最終的においたのは3億円という額でした。つまり3億円を半年で実施するという事業が承認されたわけでありまして。ただ、国に計画は出しましたが、それを実際に遂行するのはなかなか大変で、松本先生、山口先生、私で分担し、若い人を含めて必死で取り組みました。もちろん、メディアセンターが加わって初めて成り立つ事業でしたので、メディアセンターの宮内文隆さんにも加わっていただきました。メディアセンターの協力がなければ達成できなかったと思います。

社会情報教育研究センターという名称ですが、これは補助金の申請書作成の段階で候補に上り、そこに書き込まれていました。社会学部からの社会調査センター設置要望という話がありましたけれども、社会情報教育研究センターという形で少し範囲を広げたセンターにしようということになり、申請が進められました。それで、この社会情報教育研究センターとしての採用人事を具体的に進めることが8月10日に決まりまして、そのあとメディアセンターが1つのチーム、あとわれわれ3つのチームがそれぞれ事業を進めるということでテーマごとに事業にまい進することになりました。

センター設立自体については、10月29日の部長会に「社会情報教育研究センターの設置について」という提案を出しまして、ここで承認されています。ですから、私が社会調査センター構想の話を受けて、半年でセンター構想が部長会承認を得るという、かなりのスピードで話を進めることになったわけでありまして。それが可能となったのは、社会調査、統計教育、公的統計利用の三者のスクラムを作り上げていったからであると思っています。

センター設立の承認を得ると同時に、先ほどの補助金の事業も進めるわけですので、この半年間は何をやっているのかよくわからないぐらい設置準備室の皆さんはバリバリと仕事をしていた記憶があります。それで、翌年2010年の2月ぐらいには事業のめどが付きまして、補助金の事業報告書をまとめることになりました。

そうすると、実際にセンターを立ち上げる段階に入るわけですが、普通は4月発足にするのですけれどCSIは3月発足にしました。これは新しい運営方式を見込んでの考え方で、このセンターが教学それから研究支援組織になりますので、学部長がセンター長に就任されるべきだという考え方に立ったものでした。そのためには3月に立ち上げて、実務的な準備を進めて、それで4月に晴れてセンター長を迎え入れるのが良いのではないかという

ことで、3月1日からセンター規程が発効するよう2月の部長会で承認を得ました。最初はそうのように始まったと記憶しております。いろんな立場から当時のことを思い出されると思いますので、みなさんからも補足していただくか、あるいは訂正いただけるとありがたいと思います。以上です。

水上：ありがとうございます。3つのチームは、それぞれ先生方が分担されたということでしょうか。

山口：分担です。まず、申請書のところで、3つそういう事業をやるということを書いて、それぞれ誰と誰がやるみたいなことは書いて…。事業責任者みたいな感じでしたね。

水上：その時点でこの3つの柱ができあがっていて、どのように進めていくか、それぞれが計画したのでしょうか。松本先生からその当時のことなどがええますか。

松本：はい。最初のアイデアは私の記憶だと2006年の末ぐらいには、頭の中にはありました。これがきっかけになっています。2006年、着任したばかりで暇だと思われたようで、大学院の改革についてのワーキンググループに入れられました。「もしかしてこれは大学院のリストラの話か…」と思って警戒して行ったら、そうじゃなくて独立研究科に博士課程を乗せるかどうかという検討をしていて、そこでは各大学院の状況なんかもお互いに出し合って議論しました。そのときに大学院ではないけれども、社会調査がらみの研究スタイルというのはどこでもやっているわけで、これを支援するセンターがあればいいなということを書いたわけなんです。そのベースには、私が翻訳をやったクロード・フィッシャーという都市社会学者の『友人のあいだで暮らす』という本がありました。その中に「パークレーのそ



松本 康

ういうセンターの支援を受けた」とか、「データはICPSR (Inter-university Consortium for Political and Social Research) のアーカイブにあるので誰でもアクセスできます」ということが書いてあって、アメリカはそういう仕組みになっているんだと思ったわけです。それで、そういうのがあるといいなということを当時の社会学部長の木下康仁先生にちらっと話していたようなんです。翌2007年度はグローバル都市研究所の設立という話があったので、そちらのほうに注力してしまっていて、私自身はそのことはすっかり忘れていたんですが、2008年度になってから木下先生から「あの話、ペーパーつくってくれないか」という話をされ、社会調査を実施したり、受託したり、もちろん社会調査教育の支援をしたりといったような枠組みのセンター構想というのを書いて木下学部長に渡しました。そのあと部長会でどういう話になったのかはわからないんですが、菊地先生のところにつながって、「この検討チームができるから入ってくれ」といわれて加わりました。そこから先は菊地先生がおっしゃったように検討している間に補助金の話が出てきて、何しろあのとき「4億円程度以上」という表現だったんです。この「程度以上」って一体何だって話だったんですけど、い

ずれにせよ、これだけ積むっていうのは大変だねということで、もうバンと計画を膨らませて。そのときはまだ RUDA (Rikkyo University Data Archive ; ルーダ) という名前は考えていなかったんですが、データアーカイブについては先ほども言ったように、ICPSR にデータが預けてありますよという記述がこの本の中にあったり、札幌学院大学が日本で初めてデータアーカイブを設立したときに私も関わっていて、しばらくは運営委員会みたいな形で入っていたものですから、その重要性は知っていました。こういう予算がドンと付いてどうしようと思うときこそ、こういう話を突っ込んでおくのがよいのではないかと考えて、データアーカイブの設立を思いつきました。技術的にはむしろメディアセンターの方々が Dspace だったらどうかというようなアイデアを出してくださって、そのところは私にはわからないので、実質的にはメディアセンターを中心に Dspace のカスタマイズをやっていただいたというのが実情です。だから、うまくいくときっていろんな要因がまとまって、追い風になってグリーンと進むんだろうなと思います。

この事業は一種のばらまき施策みたいなもので、麻生内閣末期の施策だったので、当然、各大学がこの際ということで手を上げていて、最終的には1億削って3億円ぐらいということになったんですが、1億減ってもまだ3億もあつたんです。お金を使うのはメディアセンターの人たちが得意なので(笑)、そちらに任せて、データアーカイブをつくるとか、調査センターがらみで海外出張をするとか、おいしいところだけ我々にはいただきました。私はシカゴ大学の NORC (National Opinion Research Center) と、それからミシガン大学 SRC (Survey Research Center) に行きました。やっぱり皆さん、苦勞されているんです。だから、我々がこれから立ち上げると言うときすごく支援的で、ぜひ頑張れというようなことでいろいろ勉強させていただきました。長くなりますので、このぐらいにしておきますけど。それが初期の状況かなと思います。

水上 : RUDA の公開は 2011 年 4 月ですが、2010 年には既にデータアーカイブの話が進められていて、翌年 4 月の公開となったのですね。

松本 : 公開するためにはデータが必要で、私の手持ちのデータがいろいろありましたので、それを提供しました。データクリーニングをやって公開できる状態にするまで当然時間がかかるわけです。ですので、もう補助金事業のときにデータクリーニングは始めていました。その一方で、Dspace のカスタマイズが動いていて、これをドッキングさせてトラブルなく動く状態に確認するまでに約1年かかりました。学内公開は10月にやっていて、トラブルも特になかったので、3月から正式に公開しました。

水上 : 山口先生は、この辺りについてどのように関わられたのでしょうか。



山口 和範

山口：そうですね。自分は統計教育ということで、ずっと関わってきました。2000年ぐらいから科研費を得て、統計教育の研究とネットワークづくりをしていました。そこで、eラーニングコンテンツの開発や概念をわかりやすく説明する教材開発とかをやっていて、立教大学の中で何か実践できないか考えていたんです。2005年から教務部長をやっていた関係でずっと部長会に出ていて、この2010年のときは1年だけ全カリ（全学共通カリキュラム運営センター）部長だったんです。で、「本当は全カリでこういうのって展開しないといけないよね」と思いながら、ちょうどこのセンター構想の話が出てきました。また、社会調査士の資格をとるための科目を全学展開できないのかという考えが一方でありました。社会調査士が2004年ぐらいにスタートしていると思

います。経営学部は2006年にできているんですけども、産業関係学科のときには、すでに社会調査士制度を入れていて、経営学部ができるときも社会調査士は残すということで動いていました。この資格科目の認定は、ある意味、質保証ということも念頭にありました。シラバスを関係の協会で審査されるという経験が初めてだったのもあって、教育の質保証を強く意識したことを思い出します。2004年から、あのときは社会学部だけでスタートしたんですよね。経済はあとからでしたよね。そのときに社会学部は3学科とも同時スタートで、経営学部をつくるときにも、これは残したほうがいいと。実は社会調査士の資格が入ったとたんに、関連科目の受講生がかなり増えたんです。特に調査実習の受講生が増えすぎて大変なことになったんですけど。さらに社会調査士の科目を全学展開できたらいいよなというところにつながっています。また、2006年が最初なんですけど、海外からいわゆる統計教育の専門家の人と呼ばれていたこともあって、日本での統計教育の改革をできたらと思案していたところでこのセンターの話になったんです。社会調査だけでなく、統計の利活用も含めて広く、統計教育をどうすべきかを議論して発信していく場になればと。ただ、eラーニングの準備については実は半年間もなかったです。その中でつくらないといけないので、大変でした。統計部分のところは前からコンテンツをつくっていたので、それを生かしながらという感じです。いま、意外と海外の人はCSIのことを結構知っています。実は海外の大御所をずっと呼び続けていたことと、CSIをつくる前に、松本先生たちとは別に、自分たちはUCLAとミネソタ、それからイギリスの王立統計学会が統計教育センターを持っていたんで、そのセンターのほうにも行って、ずっとつながりを持って、その後もいろいろと活動してきました。教材もいろいろと提供してもらったりして。

水上：現在、CSIホームページの英語版はあまり充実していないので、正確に発信できるように海外とのリンクを整理する必要があるそうですね。全学的に統計教育を展開していくには、どのようなステップがあったのですか。

山口：もともと全カリに新しい科目をまとめて大量に入れるのは難しかったんですが、そのときは自分が全カリ部長だったので、新規科目として入れてしまいました。基本的に今eラ

ーニングは7科目ですよね。最初は4科目からスタートしましたが、その辺は基本的には全学共通で、なおかつキャンパスを越えてeラーニングとして提供するというので、全学展開していきました。その後、「多変量解析入門」が増えて、今、英語の科目が増えて、eラーニング自体は少しずつ増えていっています。展開としては、たまたま自分が全カリ部長だったので進めやすかったことと、あと、その後青木昇先生が自分のあとに全カリ部長になられて、必要性をご理解いただいて、そこを引き継いでもらったことも大きかったと思います。他の様々な、例えばリーダーシップとかも全カリで展開しているんですけども、そのような科目群とは違ってCSIの提供科目は内枠なんです。

	学部	科目名	
1	全学共通科目	社会調査入門	オンデマンド科目
2	全学共通科目	データ分析入門	オンデマンド科目
3	全学共通科目	社会調査の技法	オンデマンド科目
4	全学共通科目	データの科学	オンデマンド科目
5	全学共通科目	多変量解析入門	オンデマンド科目
6	全学共通科目	統計情報で社会・経済を診断する	
7	全学共通科目	景気・格差問題と統計情報	
8	全学共通科目	景気・格差問題と統計情報	
9	全学共通科目	Introduction to Statistics 1	オンデマンド科目・英語科目
10	全学共通科目	Introduction to Statistics 2	オンデマンド科目・英語科目
11	社会学部	社会調査法 2	
12	経済学部	調査実習	経営学部「社会調査実習」と併置
13	経済学部	統計調査論 1	
14	経済学部	統計調査論 2	
15	法学部	社会調査法	
16	法学部	社会調査法	
17	ビジネスデザイン研究科	統計学基礎 1	法学研究科「統計学特論」と併置
18	ビジネスデザイン研究科	統計学基礎 2	法学研究科「社会調査特論」と併置
19	ビジネスデザイン研究科	統計学 1	
20	ビジネスデザイン研究科	統計学 2	
21	観光学部	データ情報処理	
22	コミュニティ福祉学部	社会調査法	
23	コミュニティ福祉学部	データ分析法	
24	コミュニティ福祉学部	リサーチ方法論 2	
25	現代心理学部	社会調査概論	
26	現代心理学部	社会調査設計法	
27	理学部	生物統計学	

図表1：2019年度 社会情報教育研究センターから各学部へ提供している科目一覧

水上: これらの科目はすべて全カリの内枠で提供してきました。実際に科目の担当者は CSI スタッフですが、カリキュラムの枠としては全カリの中にある。当初から全学的に展開していく目的があったわけですね。

山口: ここで言っているいかちちょっとわからないんですが、補助金の本来の目的は研究基盤整備に置かれていました。だから、学部教育というよりは、先ほど松本先生から話があった大学院のところを本当は目指していたという形で調書に書きこんでいます。大学院についてもいろいろとチャレンジはしています。その後も、今、学部生に展開している科目は一応全員、修了要件の単位にはなりませんけど、大学院生は全員受講できるようになっています。各自の研究に必要となる、調査や統計の知識を得るためです。



水上 徹男

水上: 特定の研究科に独自に提供している大学院科目もいくつかありますが。

山口: たぶんそれは担当として助教が出ていっているという感じですよ。補助金はいわゆる「研究基盤整備」だったと思うので。どちらかという和社会調査とか、研究を進める上での調査の支援をするということで。CSI の中では、いわゆるコンサルとかが結構重要な業務として大きいですね。

II CSI が学内・学外の中でどのように認知度を高めてきたか

水上: CSI が学内外でのプレゼンスを高めてきた点について、1 つは今、山口先生からお話がありましたが、海外の研究者を招聘して英語圏に向けた発信があげられると思います。学内では、特に全カリを中心に、あるいは各学部でも、開設の初期段階で認知されていたのでしょうか。

菊地: 2009 年 6 月 18 日の部長会資料に書いてあるのですが、社会調査士の資格について推奨しているのは 3 学部となっています。産業関係学科の時代からの経営学部とそれから社会学部。社会学部ではまさにカリキュラムの根幹に組み入れられていたと思います。経済学部はそこまでいっていませんでしたが、経済学部も推奨 3 学部の中に入れてしまいました。ですから、後追いでそういうカリキュラムを充実させるということで進めました。あとは検討中が 2 学部という、最初はここから始まりました。ですから、全カリでの展開はありますけども、やはり学部での普及がないと難しい状況でした。そうならないかと進まないだろうということで、経済学部も協力することになりました。その後、新座キャンパスの 3 学部もカリキュラムを導入しました。全カリへの展開と学部への広がり、この両面を追わないといけなかったのが当時の状況でした。

水上：社会学部は1958年にできたとき、産業関係学科の先生が中心的な役割を果たしていて、今で言う社会調査士関連のような科目はあったと思います。

山口：実は産業関係学科の前に、産業関係研究所があって、社会学部設立と同時になんです。あそこって調査をかなりやっていたので、そういう意味でいうと当時からありますね。

水上：調査に関する教育の伝統があって、経営学部ができたのが2006年ですね。そのときに、その産業関係学科の流れに乗って経営学部ができているから、この調査関係は当然経営学部も引き継いでいて。社会学部は社会学部で、社会学科が中心になって調査科目は必修として設置されていました。2002年に現代文化学科ができて、調査科目はさらに強調されたと思います。

松本：ただ、社会調査士科目となると、ちゃんとA科目からG科目までワンセットそろえなくてはいけなくて、それ自体は社会学部といえども結構負担といえば負担なんです。それを他学部に広げる場合にも、やりたいといっても科目を用意して人を張りつけなきゃいけないのが非常にネックかなと思っていました。そういう点ではセンターがあって、そこが支援を出せると非常にやりやすくなるという頭はあったんです。逆にいうと調査士科目に引きずられると各学部の人事がゆがんでしまうんじゃないかと思ったわけです。それを避けるために、理想をいえば社会調査センターが、そういうスタッフを抱えていて、授業を全部出前でできるのが1番なんだけど、あいにくこういうセンターだとか、CSIが専任の教員を持つことができないので、そこまではいかない形で講義に協力するという段階に来ている感じですね。

山口：日本の大学ではもうすぐ、学位プログラムという仕組みが広がっていくと思います。今は、教員が所属している学部の中で、その学部プログラムができあがっているの、設置の態勢が変われば、たぶん変わるんじゃないかと思います。例えばそれぞれの、こういう調査とか統計とかのセンターがあって、全学はそこからいくつか取ってきて、学位プログラムとしてその学位プログラムを作れば、そちらのほうが効率的に、大量のカリキュラムが作れます。時代のニーズに合ったこともできるし、変えないところは変えないままにしておける。将来的にはそうなると思います。今、筑波大学はたしか大学院は学位プログラム化しています。所属と全く関係なくていいというふうになっている。将来的にはこういうセンターがたくさんできていくんじゃないかと思います。狙いはそこにあったんですよ。立教みたいに大きくない大学ですべての学部社会調査や統計の専門家を置けるかというは無理なんで、それを共通化するみたいな意識というのが結構ありました。

水上：部分的にはできているところもありますが、学部内だけで全体のプログラムの展開はできないので、全学の拠点があった方が効率化できると思います。それでは内外の認知度を高めるために、何か施策はあったのでしょうか。

山口：1番は対外的にいろんな連携を組みながらですよ。

松本: 対外的にはむしろよく知られていて、うらやましがられたり、よくやっていると言われます。むしろ学内の他部局、特に事務の部門が、CSI が何をやっているか、本当にはよく知らない。なので、日本学術振興会の補助金（人文学・社会科学データインフラストラクチャー構築推進事業）を見逃して、立候補できなかった。本当うってつけだったんですよね。もちろん出して競争で負ければそれは仕方がないけれども、スルーしてしまったというのは非常に痛かったです。

水上: その件は総長室に伝えにいきました。すごく残念でしたので、こういうことがあってはいけないと。外部資金を取って活発に活動をしていくのは、今までもこれからも重要課題なので、この点は私も注意しなければ、と思っています。先ほど山口先生がおっしゃった海外への発信はそのまま継続できているのでしょうか。

山口: 発信よりは、どちらかというと、海外の人たちとつながって、きちっと立教大学の中で成果を出していくことですよね。CSI 側が海外に対して今、貢献できるものがどのくらいあるかということ、そんなに多くはないので。今の統計教育についての新しい研究とか、先端のところとちゃんとつながって教育の質を高めていく作業のほうが大きいと思います。今までに招聘した海外研究者は、実は統計や統計教育の分野で国際的に活躍している人が多くて。例えば、2006 年のときに統計教育で呼んだ Utts さんは、2015 年にアメリカ統計協会の会長になりましたし、今の ISI (International Statistical Institute) の会長も実は呼んだことがあるんです。ISI の会長というのは、統計でいうと一種のトップみたいなものなので、そういう人たちが立教に来ていろいろとしゃべってくれているんです。統計の大切さをどうやって伝えればよいかみたいなことを説明してくれていたんです。そういう人たちが立教に来て、立教にこういうセンターがあるということを知ってくれています。あと、早稲田のデータサイエンスセンターも設立前に CSI を訪問しています。

Ⅲ 学内外の教育・研究上の課題にどのように対応してきたのか

水上: 学内の教育に関しては、全学的な展開、学部の展開は全カリなどで進めています。教育の充実と研究上の貢献で、特にわかりやすいところがありますか。RUDA は、学術的な貢献ですが、それ以外にも海外から研究者を呼んで来て、関連する情報提供をしていただくのも研究上の貢献には違いありませんが。

松本: 当初、調査を受託して展開するアイデアがあったんです。それはある意味お金を稼ぐということでもあるんですけど、結果的に菊地先生のところで自治体からの受託でいくつかプロジェクトを展開したのは意味があることかなと思っています。

菊地: そうですね、2022 年もすでに予定があるとか、継続的に受けながら展開はしています。政府統計部会の場合は、公的統計をどういうふうに教育や研究に生かすかが問題意識としてあります。国の統計法が 2007 年に全部改正されて、「行政のための統計から、社会の

情報基盤としての統計」というように公的統計の位置付けが変えられ、教育や研究面でも使えるようにし、それから地方自治体とか国の行政でもいっそう使えるようにするというのが大きな流れとしてありました。

それで公的統計を学習するコンテンツを作ろうということで、政府統計部会ではたくさんコンテンツ開発をして、学内に提供してきました。大きな広がりにはなっていないかもしれませんが、そういう取組みを学外に向けて広報することは結構やりまして、内閣府の下にある統計委員会で立教の取組みを報告し、これからは大学でもそういうことをやらないといけないということで知られるようになってきました。

それが評価されて2013年に日本統計協会から統計活動奨励賞をいただきました。それで、もっと頑張りたいということになり、その後、自治体へデータ分析に協力する中で、自治体独自の事業所調査について、愛媛県の松山市であるとか、東温市だとか、一昨年は福岡県の田川市とか、割と切れ目なく依頼が来ています。社会調査の力といいますか、調査の設計から集計分析をして行政の施策に生かしていくという、この一連の流れを応援するところ、ところが少し評価されてきているのかなと思っています。

それで、こうした調査に関連して、この間力を入れているのは、独自調査のデータと経済センサスとか国勢調査などのセンサスデータを個票ベースでマッチングさせて、さらに分析を深める取り組みです。もちろんそのためには自治体の方で国に公的統計の二次利用申請（統計法33条）をしなければならないのですが、その申請の応援もしています。

この点、東温市の産業振興でそのモデルケースが作られてきたかなと思っています。自治体では多くの場合5年とか10年の計画を立てて数値目標まで立てないといけない。そこにデータ分析をした結果を生かして計画立案をする、数値目標を含む行動指針とかを作るようにする、そういうモデルケースを各地の自治体や国の機関に紹介し、こういうやり方もあるのでぜひ頑張らしましょうと、そういう活動を今、対外的に政府統計部会でやっています。

山口:自分が周りから聞いたときに、統計調査員プロジェクトはかなり高い評価を受けていました。今は実際の調査員の、特に国とか自治体の調査を担っている人たちがかなり高齢化していますよね。一方で、若い人たちが調査員とか調査の経験がないですし、授業で社会調査実習をとっていたとしても、なるべく具体的な調査に行くようになっていると思うんですが、結局部分的にしか経験できない。一連の調査手続きを、一種のインターンシップに近い形で実現できたのは大きかったですね。

菊地:そうですね。調査員の学生は臨時公務員になるわけですね。私が退職の年に東京都から依頼が来て、CSIと、それからキャリア



センターですね。これがやはり大きかったと思います。学生のキャリア形成の中にそれを位置付けていくという形で取り組めたのがよかったです。退職後、櫻本先生が着任され、引き継いで、立派にやってくれまして。今、山口先生が言われたように対外的にはかなり評価されてきて、あの経験を話して欲しいというのが、櫻本先生のところにも来ていますし、私のところにも今、来ている状態です。

山口：いろんな自治体で、大学と協力して統計調査員のプロジェクトをやりたいということが、かなり検討されているみたいです。

菊地：そうですね。ちゃんと広まっていくとよいと思います。

水上：こちらの窓口を明確にして、ニーズを受けられる体制があることが重要ですね。

山口：あと、調査の受託に関して言うと、CSI の最初のモデルになっていたのは、たぶん松本先生がおっしゃった ICPSR の…ミシガンはかなりの部分の調査委託というか、研究の調査委託は受けているはずでもんね。あそこはすごい規模ですよ。

松本：あれはすごいです。調査会社みたいな状態で、iPad みたいなものを持って行って、その場で入れながら即時集計していく体制ができていたり、それから電話調査のためのボックスが鍵のかかった部屋に設置されていたり、これにはびっくりしました。

山口：大学の中でもんね。

松本：そう。

水上：CSI で調査を受託していくには、ある程度制約がありますが…。

松本：人手と枠組みですよ。要するに稼げる体制がもしあったとして、それで自前のベンチャーじゃないけど展開できるような枠組みはないし。

山口：ミシガンについては、アメリカの調査に関する研究費のかかなりの部分はそこで消費されていると聞いたことがあって。そういう意味でいうと、信頼されるし、そこへの調査委託であれば研究としても問題ないレベルということで、逆に個人でやるよりはそこに委託したほうが質として高いということですよ。

松本：アメリカだから民間ベースかと思ったら、意外にそうでもなくて。かなり公的資金が入っている印象を受けました。そういう点でいうと、ヨーロッパ、アメリカのインフラレベルに比べると日本は何もないに等しいわけで、ようやく一部の人たちがその問題に気付き始めたというのが現状なんです。だから、独自に学生を持っていないから全然仕事してないと思われるかもしれないけれども、そういう意味では図書館だってメディアセンターだっ

てそうで、これは大学にとって必要なインフラですよね。図書館は設置基準に入っているから作らないわけにはいかないんだけど、メディアセンターはデファクトですよね。いずれ、やっぱり社会調査がらみとか、統計がらみとかのデータサイエンスのインフラストラクチャーというのは、どの大学にもないと成り立たないインフラになっていくはずなので、そこまでどうやって頑張ってつないでいくかというところです。だんだん、その理解は進んでいるとは思いますが、非常に遅い感じがするんです。日本国内では。

IV CSIの運営を安定化させるためにどのように対応してきたのか

水上：CSIの運営会議でも前に少し話が出ましたが、「収益を上げる事業展開が果たして大丈夫なのか」ということです。例えば、この間刊行したテキスト²⁾も非常に評判がいい。こういったことで収益を上げるのは可能ですし、受託調査による収入も考えられます。収益を上げる事業をCSIが手掛けて良いのか、という点です。この辺りはいかがでしょうか。山口先生、いろいろお考えがあると思います。

山口：1つは授業料収入で支えられているので、まずは今いる学生に対して質の高い教育を提供しないといけないと思うんです。あとは研究面の成果。そして、社会貢献としての意味。社会貢献をちゃんと果たしたときに、「見返り」というと変なんですけど、そういうものを受け取れるような仕組みも必要だろうと思います。それと、たぶん今後は学び直しというか、社会人の人たちへの教育にどのくらい人と時間をさけるかを考える必要があるのかなという気がしています。今だと、外部の団体や企業から「統計検定のセミナーを実施したいので協力してください」とか、昨年から統計調査士のセミナーを有料で学外に公開して受講できるようにしていますよね。今は学外向けのセミナーは資格関係でやっているんですけど、そうじゃない形もあって、例えばムークみたいなやり方もあると思います。ただ、今は無料でムークをやれる程の余裕はないので、実際には何らかの形でお金をもらいながら社会人の学び直しに貢献する方向性で考えていければと思います。その収益というか、本当に微々たるものしかお金が入らないと逆にやるだけコストがかかってしまうと思うので、その辺のバランスは考えないといけないのかなと思いますけど。

水上：1つの方法としては外部資金でしょうか。先ほどお話がありましたアメリカでインフラ整備に政府がある程度出資するような。外部資金の獲得に関してはいかがでしょうか。

山口：外部資金については、今しばらくは、もしかすると取りやすい状況になるんじゃないかなという気がしています。今、大学内のデータサイエンスの副専攻でやっていることについての認定に関する書類を出しています。まだその審査結果がわからないんですけど、それが通っていると次の補助金にからむんじゃないかという話をされています。2年前までは5年間の連携事業でお金が入ってきていたんです。取れるところは取ればいいかなと思うんですけど、外部資金だけをあてにするのも厳しいので、あとは学内の理解を得て、学内への貢献で「CSIが必要」という認識を高めてもらうことが1番じゃないかなと思います。

菊地：それが1番重要ですね。

水上：わかりました。例えば RUDA の話にしましても、学内の認知度が把握できていないので、まずは足元、学内の状況の把握でしょうか。

山口：ただ、1番は足元をといったときに、今後検討し直したほうがいいのかなと思うのは、コンサルティングですね。特に、大学院生の。正直に言うと、11月、12月に来られるとほとんどお手上げなんですけど、秋学期に増えるじゃないですか。修士論文のコンサル。あの辺をもう少し体系的にいろんな研究科と一緒にって取り組めるといいかなと思うんですけど。

松本：それに関しては図書館と話をしたことがあって、図書館もコンサルをやっているじゃないですか。あそこは基本的に初歩的なものを受けたがっているんです。高度なものを持ち込まれると困っちゃうし、あんまり人が来ないのも困ると。で、こちらはこちらで高度なものを期待しているんだけど、割と素朴なものが持ち込まれると困るというので、お互いに連携して紹介し合うようなことをやったほうがいいんじゃないですかみたいなことでちょっと話をしてきたところなんです。だから、図書館との連携は、そういう意味では細かいところで必要になってくるかなと思っています。だんだん、CSI と図書館の間で「お互いにやっているのね」というのがわかってきたので、少しスムーズになるかなと思うんです。あと、私がもともと狙っていたのは、むしろ大学院生じゃなくて研究者そのものの。

菊地：それは外部も含めて。

松本：そうです。科研費を受けて、それで調査をやると。そうするとアイデアは自分で持っているけれども、ノウハウはわからないというところに対して CSI が協力する体制があればと。そのときに科研費から委託費の受け渡しができるといいんだけど、今はそれができませんよね。調査会社に出すのであれば支出できるけど、学内にはお金は出せないですよ。だから、やろうとすると結局共同研究者という形で名前を入れればそこで使えるけれども、そうでないとできないというんで、それは枠組みを考えながら。結局アメリカなんかだと、調査の細かいテクニカルな設計をやる専門家と、その研究のアイデアを持っている研究者とは別々なんです。アイデアを持っている人はアイデアしかないんで、調査の専門家に相談すると、「こういうアイデアだと、こういう調査設計でどうですか」って準備してくれるわけです。それが今、日本だと全部できる人がやらないといけないみたいな状況なので。それは分業体制で展開できることが重要だと思います。そのための仕組みも必要だし、そういう研究スタイルに、だんだん転換していくということで、調査の専門家と特定の実質的な研究分野の専門家がコラボできる組織体制を準備していく必要があるのかなと思っています。

山口：お金が動かせるということで言うと、外部だったら簡単に受け取れるので。内部の科研費を CSI に渡すのはできないので。あれはあれで不思議ですよ。

水上：そうしますと、内部の相談の方が受けにくくなりかねないという。

山口：いや、お金をもらおうということと言うと、です。だから、お金をもらわずにやるということで、裏からお金に来るようにという意識をちゃんとみんなが持ってくれるといいと思うんですけど。あとはたぶん、場合によっては研究計画書を書くときに、調査はCSIと…みたいなことを書いて、そこはそれで信頼できるようにCSIがなっていくと。

松本：そうですね。それが理想ですね。

V 現状の課題と今後の展望

水上：昨今のデータサイエンスやAIの状況を踏まえて、CSIを含めた学内組織の再編が求められているのかもしれませんが、そのあたりはどうでしょうか？

山口：キーワードとしてデータサイエンスのほうがいいのか、今みたいにスタティスティクス・アンド・インフォメーションのほうがいいのか、「社会情報」といったときの意味合いとか、結構日本語と英語で違うんですけど、両方含めて広くとれるようにはつくっているんです。

菊地：スタートのときの社会調査と統計教育、それと政府統計の活用、この3つの組み合わせがすごくよくて、学内的な議論を進めていく中で理解を得やすかったです。先ほどの「全学データサイエンス構想チーム」にもCSIから委員として入られていると聞いております。この構想チームがスタートするもともとの問題意識のなかにデータサイエンス教育を全学部で展開すると補助金のカウントが上がるという発想もあるかと思いますが、大学の教学サイドからみると、社会に出てから必要になるいろんなデータリテラシーのところで、ある程度提供できている学部とそうでない学部がある。そうでない学部が少なからずあるのはまずいのではないかとということで、全学に提供できるようなことを考えないといけないのではというのが問題意識の根底にあるようです。たしかにデータリテラシーの教育はまだ限られている気がします。学部での展開、これをやはり広げる方策をぜひ考えていただきたいと思っています。

先日、日本学会が「公的統計問題を学術の視点から考える」というテーマでシンポジウムを開きました。まさに統計不正問題の学術への影響を考えるということです。私も出席



しましたが、これまでは公的統計に関わる日本統計学会や社会調査協会の皆さんとで一緒に議論することがあまりなかったように思います。今回の場合はさすがに統計不正の問題があれだけ国会で議論になって、このままではまずいだろう、研究面に少なからず影響を及ぼしてくるだろうと、社会調査の人と統計に関わる人が一緒になって議論をしました。

実は、そういうことがこれからますます必要になり、協力して進んでいかざるを得なくなると思います。行政の人に社会調査の力をきちんと付けてやればよいといった言葉が飛び交ったりもしていましたが、そうになっていくとこのセンターがスタートしたときの組み合わせ、これが社会的にも一層必要とされてくるのではないかと思います。CSI センターの名前の変更の議論が出てくるかもしれませんが、現在の名前の意味をしっかりと考えてさらに進めていただけるとよいと思っています。CSI は社会の動きにますます合致していくのではないかと思います。

松本：先ほど山口先生のおっしゃった統計調査員のインターンシップにしても、1 番のデータをとるという足腰のところがすごく弱くなってきていることへの対応ですよ。データをとる部分がいい加減だと、いくら分析を高度化したところで話にならないわけで、その重要性というのがだんだんと認識されるようになってきたかなとは思っています。

山口：特に今の AI がデータに基づく AI じゃないですか。そのときに観察とか記録というときに、何に気を付けないといけないかというのは、調査のリテラシーというのがちゃんとないと非常に危険だと思うんです。

松本：ただ、ビッグデータでないと分析できない部分というのは確かにあって。ビッグデータは絶対数は多いわけだから、どんな小さな動きでも、ちゃんとそれはそれで捉えられる点では意義があるかなと思います。ただ、AI は人間にわかるように説明はしてくれませんか。間に人間が入って解釈してやらないといけない。そのところはデータサイエンスが、例えば社会現象なら社会現象を捉えるときに、どういうふうに機械と人間の頭との媒介をしてやるかという。そこではかなり貢献できると思います。あと、社会調査部会的な課題でいうと、RUDA でデータはアーカイブしているんですけど、やっぱり自分たちでデータをつくらないといけないと思っているんです。なかなかその体制はできないけれども、やっぱりある種の General Social Survey みたいなものを企画して、毎年とは言わないけど定期的にそれをやってはアーカイブしてって、それは割と研究教育に使えるスタンダードなものですよ、質のいいデータですよというのを提供できると理想的だなと思う。

山口：学内的に CSI に期待されているのが IR での貢献なんです。それはかなり前からで、今は新しくいろんな検討が進んでいるので。そのときに、人も含めて。

松本：人的には貢献しているんですけどね。今でも。

水上：IR 自体も調査ですからね。

山口：実は今、経営学部は学生も巻き込んで、外部資金による分析チームがあって、追跡調査とか、リーダーシッププログラムについての評価を行っています。学部としては、外部資金が終わっても、その調査を継続して、教育強化につなげていく方針です。

水上：IR もデータアーカイブで使えなくはないですが、対象がすべて立教大学内で進めていますね。

山口：学生の情報なので、外には出せないですね。

水上：そうですね。今は学内なのでオープンにできませんが、あの種の調査ですとか、松本先生ご指摘のように何らかの社会調査を実施してアーカイブで出していくという。

山口：日本社会学会でやっている SSM 調査³⁾、あそこの母体はどこなんですか。

松本：あいまいなんです。SSM は伝統的に、研究者が SSM グループを組織してやっている。現在はデータが整理されて、東大の社研のアーカイブで公開されています。それから家族社会学会は、学会として「全国家族調査」というのをやっています。

水上：学会レベルではないので、CSI の余力ってそれほどありませんね。

山口：逆にいうと、科研費を申請して取って、継続して調査やデータアーカイブの運営をやっていくのは、本当は考えないといけないのかもしれないですね。

菊地：若い人を育てる。育ててもらいたいですね。

松本：あとは、なんかの機会にこういう補助金がまた来る可能性があるんで、そのときにチャンスですね。常に夢を持ってね。もしチャンスがあったらこういうことをやりたいということを持っていないといけない。持っていればドカンと来たときに、それだったらこれに突っ込もうとなるので。

水上：ソーシャルサーベイとか、菊地先生が行った自治体との連携などは、できそうですね。特に今、豊島区は活気づいているので、地の利もありますし、立教大学と豊島区でプロジェクトは組めそうな気がします。

山口：そういう意味だと、本当はやっておいたほうがよかったなと思うのが、いくつかの、例えば自治体だったりセンターだったりとの包括協定みたいなことですね。

水上：つくるべきですね。

山口：本来であれば、和歌山にある統計データ利活用センターとは連携をお願いした方がい

いですよ。早めに。あと、日本統計協会とかも可能であれば。

水上: 人材確保も問題になりますが、調査設計とか、学内での認知度を高めることについてはニーズがもともとあるわけですから、ニーズがあるのはわかっていて、幅広くアクセスがあり、こちら側はアーカイブを充実させるためにソーシャルサーベイをやりますとか、自治体と連携しますとか、これら全部を、山口先生と私では到底できない。

山口: いやいや、無理ですよ。今でもポス特的に、学内的にいうとぜいたくさせてもらっているセンターだとは思いますが、本音を言えばもうちょっと増やしていきたいところですよ。あとは、学部プログラムも含めて一定程度のところを全学共通で担うので、教員として。

水上: もう少し採用させていただく。

山口: 自分は専任教員として CSI に移ってもいいと思っているんですけど (笑)、ずっと。

水上: それでしたら交渉のしがいがありますね。専属も CSI は準備できる (笑)

松本: だけど、専属は必要ですよ。

山口: 絶対、そう思うんです。だから、専属で何人かいて、あとはうちでいうと学部には属さない教員を、今度ようやく言語のほうでそういうセンターができるので、定年前に移っていいですよ。経営学部にも怒られそうなんだけど (笑)

水上: 定年前に移っても、定年延長でまだ何年もあるから。

山口: そんな長くはやりませんよ。たぶんそういう形で専任がいないと。

水上: そうなりますと展開がかなり変わりますね。

山口: 自分はいろんな学部にも教えに行き全然構わない。教えていいと言われてれば。

水上: 学部スタッフとさらに連携した組織になればいいですよ。兼担という意味ではなくて、運営を生かした学部、今、我々が行っているように学部組織との連携もあって、CSI 専任の先生も存在する組織ですね。実際に今まで、授業提供のほとんどは助教の先生方が担当してきましたが、それだけでなくさらに充実したシステムに向けて。

菊地: 大学では教学条件というものを作りましたよね。あれは私も最初にいろいろ関わったりしたんですけど、それから管轄人件費という、この枠組みは、学部関係ではすごくきめ細かくできています。しかし、そこから外れているセンター、機構がものすごく多くなって、

さすがに大学としても無視できなくなっているのではないかと思います。そういうところの収入の割り当て方をどうしていくのかを明確にすることが大事で、CSI はもっと教学条件の中に入るような仕組みで位置付けられるべきかと思います。今は完全に外で、この枠内にとことをぜひ目指していただければと思います。

山口：だと思います。大学全体として考えていかないといけないのが、CSI が、例えば学部のようにいたり、全カリで科目を持っているということも含めて、実際に各学部からお金をもらって教学条件、いわゆる学部管轄人件費的なものを、CSI が持って、その中でちゃんと各学部と交渉しながら、「おたくのこれだけのコマを持つので、これだけの人件費をくれ」と言って人を何年間か計画立ててみたい。例えば今は学部管轄人件費があるので、かなり自由に計画を立てていろんな教育をやってみたりできるんですね。グローバル教育センターもそうだし、いろんなセンターが、たぶんそういう仕組みにならないといけないんです。

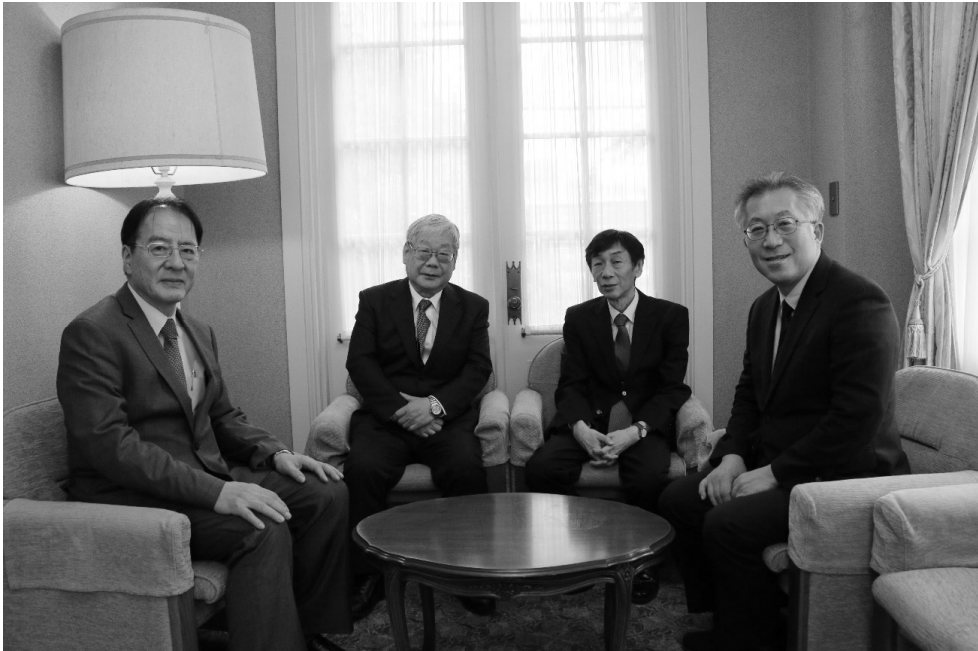
菊地：そうですね。そう思います。

山口：そういう意識がずっと頭の中にあっていろいろな仕組みをつくってきて、CSI だけが全カリの内枠なんです。他はどうしても内枠にできなくて、グローバル教育センターも実は全部外なんです。サービスラーニングセンターも外なんです。いわゆる外にすると、スクラップアンドビルドにならないんです。収入は増えずに支出だけ増えるんです。バラエティは増えるけど、それだと大学全体としての経営としては成り立たないので、たぶんそこは大学全体でちゃんと考えていかないといけないと思います。言っては悪いけど、学部の立場でいうと、学部としては「これだけのものは自分たちが持っているから自分たちが使う」という考えになってしまう。そこから外には出さないよという意識があるんだけど、実際にはいろんなところに、CSI の資源を使って教えているときには、タダでもらっている状態になっていますよね。

水上：そうですね。

山口：それは変えないといけないと思っています。「CSI にはニーズがある」という認識を各学部が持ってくれれば、CSI は拡大できるし、そのためには学部に対して、ここが使えるセンターだということをアピールしていかないといけないので、ある意味、いい意味でもあるんじゃないかなと思うんですけど。

水上：CSI の活動を理解していただき、今後の展開には、発信することが重要だと思います。今望まれているのは、調査のリテラシーであるということを、明確にこちらから打ちだして、その上で各学部で認識していただくと、特に AI に注目が集まっていますが、このような時代ですからおさら、こちらのからのメッセージを正確にだしていく必要に迫られます。本日は、お忙しいところ本当にありがとうございました。



注

- 1) 松本 (2012) を参照.
- 2) 櫻本・濱本・西林 (2019) を参照.
- 3) SSM 調査とは、日本において 1955 年から 10 年ごとに実施されている「社会階層と社会移動 (Social Stratification and Social Mobility)」に関する大規模な全国調査である。第 1 回の調査は日本社会学会を基盤として実施されたが、第 2 回目以降は社会学者を中心にした研究組織がその都度軽視されて調査研究が行われてきた。(一般社団法人 社会調査協会編 2014: 700)

参考文献

- 一般社団法人 社会調査協会編, 2014, 『社会調査事典』丸善出版.
- 松本康, 2012, 「立教大学データ・アーカイブ RUDA の始動 (特集 データ・アーカイブと二次分析の最前線) -- (量的調査データのアーカイブと二次分析)」『社会と調査』8: 24-30.
- 櫻本健・濱本真一・西林勝吾, 2019, 『日本の公的統計・統計調査』立教大学社会情報教育研究センター.

(24) 社会情報教育研究センター研究紀要『社会と統計』第6号

《論文》

ミクロデータ分析と公的統計データベースに関する展開 2019 年度出張報告を兼ねて

櫻本 健
西林 勝吾
濱本 真一

【要旨】 2007 年に全部改正された統計法は、2018 年に再度改正され、2019 年 5 月に施行された。法令の制約により一気に制度変更は進むことがなく、制度変更、研究者による分析が繰り返し循環する中で、少しずつ制度的枠組みが進化する。日本は北欧と比べて統計制度の利便性が低いことで知られているため、法改正と運用の改善によってこの差をいかに縮めるのかが、社会的に重要な課題となっている。

そこで本稿は統計改革の進捗を概観できるよう、政府統計部会のお出張報告を兼ね、統計法 2019 年改正後ミクロデータを用いた研究にどのようなものがあったかサーベイを行った。その結果、オンラインの具体的な事例、公益を認定した経済センサスほか母集団名簿の利用、その他公的統計と民間データを組み合わせるなど、公的統計制度を前進させられると思われる研究は、統計法改正後 8 か月を経た段階でもほとんど見られず、統計法改正前後で利便性に大きな差はみられなかった。利便性が高い主要国の二次的利用制度とは依然大きな差があるとみられ、統計法の運用面での対処が今後の課題となろう。統計法改正後、オンラインサービスは徐々に拡大すると思われる。公的統計の二次的利用制度では利便性の改善が著しく、行政データの利用環境が整えば、今後現在の研究の方向性が制度面の改善につながる可能性がある。

キーワード：公的統計の二次的利用制度、統計法改正、ミクロデータ、オンライン方式、リモートアクセス方式

I はじめに

統計法は 2007 年に全部改正された後、2018 年に再度改正され、2019 年 5 月に施行された。そうした最中に統計不正問題が起き、社会を揺るがすこととなった。2018 年 12 月、厚生労働省が所管する「毎月勤労統計調査」の調査プロセスが、適切に行われていなかったことが明らかになったのである。本稿末の参考文献にまとめた 6 つの新聞記事によると、この調査は全国 3 万ほどの事業所を対象にした調査で、都道府県を通して実施しているものである。それが、東京都において 1400 事業所を調べなければいけないところが、実際は 500 程度しか調査されていなかった他、2004 年から 2017 年まで厚生労働省が放置していたことも注目を集める要因となった。とりわけこの調査は、雇用保険の給付などに用いられていたため、専門家だけでなく世論の関心も高かった。この件は、他の公的統計にも波及し、適切に調査が行われていたのかを確認する事態となった。統計不正再発を防ぐため、内閣官房にチェック要員 30 数名を配置し、統計実務を行うための資格創設なども検討され、統計法改正に盛り込まれていない改革が数多く施行後に準備されるという、前例のない混乱につながった。

統計不正問題が世間の注目を集める中で、新統計法の改正内容の骨子はごく地味で専門性も高いため、新統計法の趣旨を理解し、統計改革を粛々と進めていくこともまた社会にとって重要となる。ここで説明しようとしている統計法改正のポイントは、総務省 (2019) や阿部 (2019) にまとまっている。両方とも統計センターウェブサイトや二次的利用制度のサ

ービスで提供されている事実上の公開情報である。阿部（2019）によると、オンサイト施設の利用を前提に調査票情報の利用の拡大が図られた。またオンサイト施設ではシンクライアント端末を用いて申請の手間が減り、手数料が減額された。またオンサイト施設では統計データと行政データを組み合わせて R、SAS といった統計ソフトを使用した高度な分析が可能で、符号票やレイアウトも公表予定のため、これまでよりもマイクロデータの利便性が大いに上がることを強調していた。

つまり、統計法改正によって期待されているのは、公的統計の二次的利用制度のうち、主に調査票情報の利便性を通じて、公的統計の利便性を大きく向上させるということである。今後は法の運用体制の整備に伴い、主に統計データと行政データとの個票レベルでの突合や連携が深化する方向に統計改革が進むということである。

マイクロデータの仕組みは、政府や公的機関同士の運用と、研究者も含めた公益が認められた広義の仕組みの 2 つに分けられる。いずれにしても北欧の国々がこの分野で最も進んでおり、そうした国では日本で主流の調査統計が衰退している一方で、行政データが統計データの中軸になっていて、政府内の統計部局ではリモートアクセス方式をベースに広範囲の推計担当者が仮想空間で情報を共有する仕組みを築いている。一方特に人口が多い国では、マイクロデータの利用制度が遅れる傾向にあり、アメリカでは日本同様オンサイト機関でのマイクロデータ利用が前提となっているため、研究者も含めてマイクロデータの利用に限界がある。日本は主要国の中で最もマイクロデータの利用が遅れた国であるため、新統計法を活用して北欧も含めた諸外国の制度に追いつくように努力している。現在は統計データに行政データは限定的にしか利用されていない。概ね現在利用されている行政データは登記簿、労働保険データ、税務データの集計値に限られる。

今後具体的には付加価値税（VAT）といった税務データをレジスターデータベースに突合して分析する研究や実務、行政データによる匿名の開発、データリモートアクセス方式に基づく行政データによる匿名データを使った統計作成・分析といった方向に統計改革が進んでいくことが望まれる。

以上統計法改正のポイントと改革の方向性をふまえ、現段階での統計行政が置かれている状況を検討したい。本稿では、旧統計法からの二次的利用制度をめぐる歴史的経緯を検討した。さらに統計法改正による利便性の向上を概観するため、法改正前後の時期を中心とし、サーベイした。統計法改正以後、統計改革の進捗を見るために①日本の調査票情報を利用した国際研究、②オンサイトの具体的事例、③公益を認定した経済センサスほか母集団名簿の利用、④その他公的統計制度を前進させられると思われる研究といったポイントを調べたが、法改正後も 8 か月を経た本稿執筆段階では大きな変化が無いことが分かった。

本稿の執筆にあたり、今年度本学社会情報教育研究センター政府統計部会は、分担して研究集会に参加し、成果をまとめることにした。実際に研究集会は多数ある中から 2 つに絞り、経済統計学会全国研究大会（東北学院大学、仙台）と統計センターが開く、共同研究集会「官民オープンデータ利活用の動向及び人材育成の取組（2019 年度）」（統計数理研究所）に参加した。それ以外は、研究会への参加や論文・研究ベースに芽づる式に調べたものを利用した。なぜこのような回りくどいサーベイをしているのかというと、日本では研究ベースでマイクロデータの分析手法が進化するのではなく、基本的にマイクロデータを扱った公開・非公開の研究会で情報が公表されることが多いことに基づく。

Ⅱではこれまでのミクロデータ利用環境の論点を整理し、法改正後の動きについてまとめる。Ⅲでは消費動向指数を中心に公的統計と民間データの組み合わせや改善に向けた取り組みを議論した¹⁾。

Ⅱ ミクロデータの利便性の向上

1. 統計センターによるサービスの変化

ここでは統計センターが提供するサービスと公的統計（公的と言っているが、現状では事実上政府統計に限定されている）の二次的利用制度について概要をまとめる。近年、国レベルでのデータ利活用促進の動きが活発化してきている。それは、内閣府の「経済財政運営と改革の基本方針 2017」などの公的な発表資料等で、EBPM（Evidence Based Policy Making）の重要性が指摘されていることをみても明らかである。国がデータに基づいて、政策立案や評価を行うこと、また研究者などといった民間での公的統計の利用促進をしていくことが求められている。その際、議論の前提にあるものはデータの整備・活用である。ここでのデータとは、主に公的統計制度で生み出された統計データを指す。

公的統計の二次的利用制度にはオンサイト方式（隔離 PC で個票利用する方式）とリモートアクセス方式（ネットワークを利用できる情報端末から個票にアクセスする方式）の 2 つがあり、後者の方がオンサイトよりも利便性が高いことが知られる。このうち、日本の現状では前者のみ認められているが、後者も運用面で実現できないということに過ぎないため、セキュリティや匿名性といった運用上必要なハードルをクリアできれば、いずれ省庁でも導入される可能性がある。新統計法において当面はオンサイト方式を中心としながら、政府統計では調査統計中心の日本の統計行政が変化し、行政データが調査データと融合することで調査から行政データへの依存が進むことが予想される。

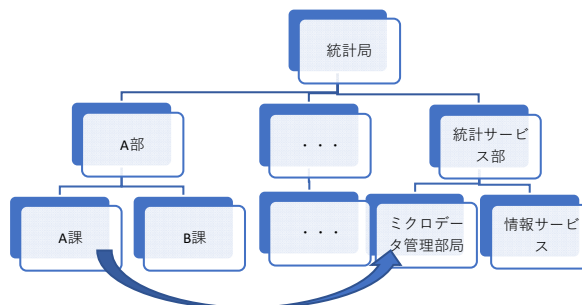
データを主に管理し、提供を行っているのが独立行政法人統計センター（以下：統計センター）である。統計センターは、簡潔に言えば国、地方自治体、研究者等の民間部門の三者をつなぐ役割を果たしている。2019 年 5 月 1 日に改正統計法が施行されたことで、当センターの役割がより重要となってきた。図表 1 はミクロデータの利用の際のサービス全体をまとめた概念図である。ミクロデータの利活用の場合、国によって制度が異なるが、日本では統計ごとに各府省庁が縦割り提供している。



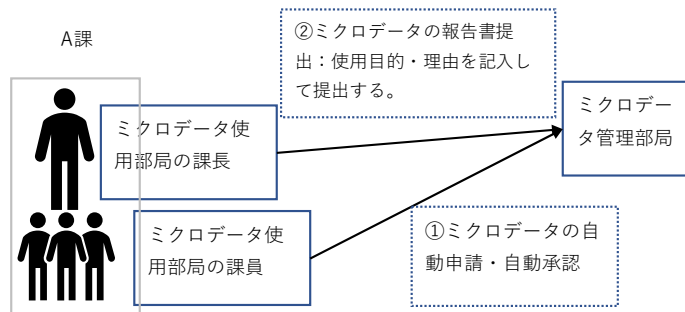
図表1 統計センターが提供する二次的利用サービス

出所：統計センターHP「公的統計のマイクロデータ利用」より引用。

筆者の一人である櫻本は2017年にカナダ統計局の推計環境を視察したが、カナダの統計システムの場合、特定の部局がマイクロデータへのアクセスの承認・監督プロセスを一元管理している。フランス、オーストラリア、ニュージーランドといった国々も同様であった。いくつかの主要国ではマイクロデータや行政データを集中的に扱う部局が存在しているが、多数のアクセスが常時生じているからリモートアクセスが必須であるし、縦割りで管理しているのは業務がパンクしてしまう。その問題を緩和するために、統計作成機関職員はデータにアクセスした時点で自動的に申請書を管理部局に提出するシステムとなっている。図表2はマイクロデータの申請イメージで、ここまでは日本も主要国もほぼ同様だが、日本の場合、紙での申請のみとなっているのに対し、主要国の多くでは電子申請である。過去にヒヤリングした限りではアメリカセンサス局はオンサイトを利用した申請権限があるが、商務省経済分析局は加工統計や分析を所管しているため、マイクロデータ利用の権限がないということであったから、一般に分散型の方が縦割りの弊害が生じやすく、利便性が低い傾向があるかもしれない。



図表2 ミクロデータの申請イメージ



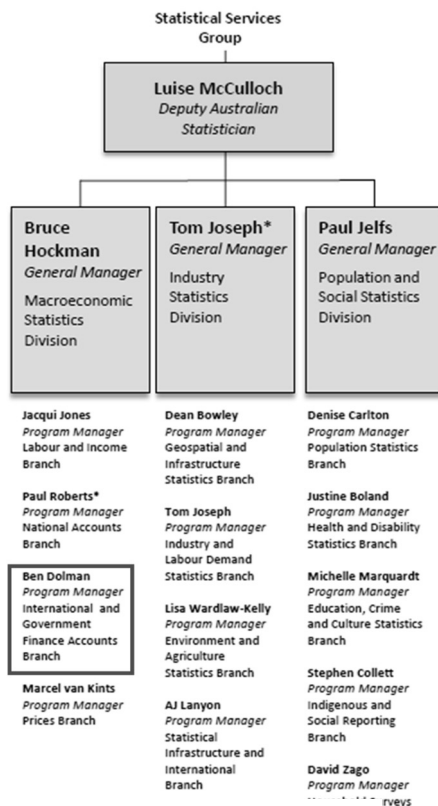
図表 3 リモートアクセス時の申請やり取り

図表 3 はカナダ統計局での申請イメージを図にまとめたものである。守秘義務があり、具体的には公表できないため、イメージで説明する。図表 2 の A 課課員が供給使用表や国民経済計算の推計といった目的でミクロデータにアクセスするとしよう。データにアクセスしてファイルをダブルクリックするだけで、ミクロデータの申請をシステムが自動で行い、リモートアクセスで接続される。必要な作業は課員レベルで自由に行い、結果がどうなったかを、課長に報告する。課長は申請に目を通し、報告書に使用目的や理由を記入する。すべて電子的に行うため、極めて利便性が高いように思われるかもしれないが、アクセスが簡単な分、それに応じた報告書を出さなければならず、期限もあることから管理者が忙しいという課題がある。

ミクロデータの管理を担当する部局があり、そこが申請や報告書の一元管理をしている。通常集中型のシステムの場合、統計局全体に情報システムのようなインフラ整備や推計マニュアルを整備する専門性の高い部局があり、ミクロデータの管理もその管轄で行われる。縦割りに管理していないという意味では帝国データバンクなどのデータが収録された RESAS でも同様となっている。日本の二次的利用制度の申請管理は依然として過渡的な手法に頼っていると言えるため、一日に申請が千件程度生じるような社会を想定して、組織の専門性を高め、いずれ総務省内や統計センターなどに集約すべきかもしれない。

統計センターには、大きく分けて提供業務の内容が三つある。第一に、「A オンサイト施設での調査票情報の利用」である。集計データのもとである公的統計の調査票情報をオンサイト施設で閲覧、利用することができる。利用の際はデータの持ち出しはできず、施設内に限って利活用することができる。分析結果は、所定の審査を受けたのちに提供を受けることが可能で、データを活用したより深い示唆に富んだ研究成果が期待できる。第二に、「B オーダーメード集計」を行うことである。オーダーメードの集計表を作成し、提供を受けることができる。これによって行政機関が発行していない統計表による分析が可能になる。そして第三に、「C 匿名データの提供」である。これは、公表されている集計データのもととなったミクロデータの提供を受け、一定期間利用できる制度のことを指す。匿名データとは、調査票情報を個人や団体が特定できないように加工して作成されたもので、一般には公開されていないが、申し出により匿名データを活用し実証分析を進めることが可能となっている。最近では、ミクロデータ利用ポータルサイトの「miripo」が誕生し、ミクロデータ利用のハードルが下がってきている。ただし、上記に挙げたデータ利用には、各手数料やデー

タの返却などといった留意すべき点も多く、それらを確認しておく必要がある。また、統計センターはデータの二次的利用サービスの相談にも乗っているため、利用に関してはアドバイスを適切に受けることができる。



図表4 組織的利用の例

出所: ABS HP 組織図より抜粋

はないため、オンラインの利便性の拡大やリモートアクセスの運用といった利便性の向上とともに、組織的ユーザーを創設して政府の能力を向上させる必要がある。行政データと二次的利用制度については、統計研究会（2018）が主要国の仕組みを詳しく解説している。

以上とは別に一般用疑似マイクロデータが開発されて、使用されるようになっている。疑似マイクロはマイクロデータを使用する際の環境に慣れるために公表されるようになったが、あくまで一部のサンプルデータに過ぎないため、依然として使い勝手の良いデータとは言えず、あまり認知も広がらず、使用されていない。

2. 匿名データ利用の詳細

2019年4月1日現在、図表5の通り、統計センターでは総務省が主管する六つの調査の匿名データの提供を行っている。

実は日本にはまだ十分に存在しているとは言えないが、国際的には第4のサービスがある。図表4は2017年時点のオーストラリア統計局の組織図の一部である（組織改編によって現在は多少変更されている）。四角で囲ったところが政府財政統計課で、財務省の行政データを基に現金収支から発生主義（修正現金収支）に転換し、政府収支を詳しく公表する政府財政統計（GFS）を推計する部局である。図のすぐ上の国民勘定課（SNA推計部局）にGFSデータを送付している部局と思われる。GFS担当部局は日本以外の多くの主要国に存在している。カナダでは概ねGFS部局だけで50名程度いる。GFS以外にも各国の組織図には税務データが匿名化し、加工統計部局に手渡す専門チームがいる。日本以外の主要国では税務データを個票データと一緒に統計局内で使用しているのが一般的となっている。過去に数回ヒヤリングした際にはデンマーク統計局では事業統計部に付加価値税（VAT）レジスターチームがおり、統計局職員が税務データを利用できるサービスを提供しているとのことであった。このようなマイクロデータを集中管理したり、組織的に行政データと突合する専門部局は現在の日本に

調査実施機関	調査名	提供年次	平成31年度 提供予定年次
総務省	国勢調査	平成12年,17年	平成22年,27年
	労働力調査	平成元年～24年	－
	住宅・土地統計調査	平成5年,10年,15年,20年,25年	－
	全国消費実態調査	平成元年,6年,11年,16年	－
	就業構造基本調査	平成4年,9年,14年,19年	－
	社会生活基本調査	調査票A(生活時間編/生活行動編) 平成3年,8年,13年,18年 調査票B(生活時間編) 平成13年,18年	－

図表 5 匿名データ提供対象一覧

出所：統計センターHP「匿名データの利用」

(<https://www.nstac.go.jp/services/anonymity.html#ano03> をもとに作成)

なお、他府省の匿名データについては「miripo」で詳細について参照できる。ここでの提供データは最新でも 5 年以上前の情報で、データの匿名化処理のため時間がかかるのは仕方のない部分もあるが、研究成果につなげるためにもより迅速なデータ提供体制の構築が望まれる。

データの利用は、申出書を提出し承諾されたのちに、依頼書と誓約書、そして手数料を納付することで可能になる。なお手数料は、事務手数料 1950 円、提供ファイル 1 ファイルごとに 4450 円、格納する媒体 CD-R100 円などと、最後に郵送料を入れた合計となる。提供を受けるデータによってファイル数が異なる。オーダーメード集計、匿名データは、統計センターに加えてサテライト機関（一橋大学、神戸大学、大学共同利用機関法人情報・システム研究機構）がサービスを提供している。

二次の利用制度では匿名データの提供機関と調査票情報の取り扱い機関が重複しているケースも一部にみられるが、基本的に別系統に属して縦割りとなっている。調査票情報は統計センター、和歌山県の統計データ利活用センター以外は全国の国立大学を中心に 11 か所にオンサイト施設が設置されている。私立大学は採算面で厳しいため、多摩大学だけがメンバーである。二次的利用制度に関しては、統計センターウェブサイト掲載の「統計センターが提供する公的統計のミクロデータ利用サービスについて」に概要が載っており、利用に際してはそれぞれの機関に相談するのが一般的である。

3. ミクロデータ活用に関する先行研究の流れ

ミクロデータ利用に関して、90 年代後半から昨今にかけて多くの研究が行われてきた。それらを時系列でみたのちに、総務省などで行われた検討の資料を概観する。

まず海外のミクロデータの状況をみたものとして、やや古い資料であるが石田（1999）や森（1999）などがある。前者は、アメリカとカナダにおいて行った調査結果をまとめたものである。アメリカでは、1963 年に 1960 年人口住宅センサスを、カナダではその 10 年後の 1973 年に 1971 年人口センサスのミクロデータファイル提供を開始した。この 2 か国の提供方法は大きく異ならず、違うのは一般使用ミクロデータに関してアメリカでは主に料金さえ払えば誰でもアクセスできるのに対して、カナダでは使用者によって提供方法が区

別されている点くらいである。提供媒体は CD-R が中心で、インターネット経由でも行われるようになってきたとある。後者の研究では、イギリスに関して触れており 1993 年から提供され始めた人口センサスデータについて、特徴や提供条件などを細かく述べている。これらの研究の意図としては、日本でのデータ提供のあり方に対する助言だと考えられる。マイクロデータ提供の法体系、制度を参考にするため、それらを伝えようという意志がみられる。同じ年の研究である、渋谷 (1999) はマイクロデータ活用のための理論と技術について、制度 (公有化) の完成と運用活用の 2 ステップに分け、課題をまとめている。今では当たり前のことではあるが、匿名化による安全性に今一度着目している。しかし、まだこの段階では日本でのマイクロデータ提供について踏み込んだ議論がなされていなかった。より体系的に現状と課題が浮き彫りになるのは 2000 年代に入ってからである。

次に森 (2004a,b) から統計法規の確認を簡単にする。「統計法」は、1947 年 (昭和 22 年) に制定され、以来その本質は変わっていない。第 2, 3 条において指定統計の定義と法的根拠が与えられ、研究当時は第 14 条には秘密保護規定、15 条で調査個票の統計作成目的以外への使用禁止を規定していた。そして、第 16 条は調査結果の速やかな公表を統計実施者に義務付けているとした。また、政府統計の二次的利用に関して同論文では先行研究を参考にしつつ、その制度運用について整理している。もともと二次的利用については「統計法」第 15 条の 2 項にあたる目的外使用という例外的な提供ルートと制度運用が行われてきた。統計上の目的とは、事前に承認を受け公示された集計方法による集計事項を集計した結果表を作成することに限定されていた。要するに統計作成機関であっても統計の二次的利用がほとんど認められていなかったのである。

細かい解釈については本論文を参考にしてもらいたい。この当時の制度運用の問題点として、主に政府機関以外からの目的外利用申請、とりわけ純粋な学術研究という目的だけでは承認の対象から除外されてきたことを挙げている。すなわち、利用目的の公益性が重視されていたのである。この点は、井出 (2004) も 1995 年以前のマイクロデータ活用状況について、「統計法」第 15 条の統計上の目的以外の使用は、総務大臣の承認を得て使用目的を公示して利用できる一方で、公益性の高さや使用者が原則公務員などといった高い基準であったため、その利用が限られていたと触れている。同井出 (2004) は現状に関して続けて、「統計情報活用フロンティアの拡大」の活動や「マイクロ統計データ活用研究会」の活動が行われていたといい、特に後者は統計調査の個票データの部分標本 (リサンプリング・データ) をマイクロデータとして提供することのユーザビリティを検証するプロジェクトを行い、それを総務省統計局も高く評価して目的外利用を承認したとある。この 2 つの研究が行われた頃は、マイクロデータをより学術研究にも活用していかなければという機運が持ち上がり始めた時期と考えられる。

その後、2007 年 (平成 19 年 5 月) に統計法が改正された。総務省によれば、社会の情報基盤としての統計、国が行う統計調査や統計データの利用促進などがこの改正のポイントとして挙げられている。山口 (2008) にも、この改正で個票データの二次的な利用に関する規定が設けられたとあり、個票データの利用にあたる課題を提示している。この研究では、先の森 (2004a,b) と井出 (2004) のように、公的統計データの二次利用の流れを確認し、それ以降の統計法の進展をまとめている。まず、「統計法」の第 15 条第 2 項の目的外使用にあたる厳しい基準を満たし、承認を得たケースはほとんどなく、大学の研究者が申請して

承認されたものは年間 10 件程度だったという。その後、特定領域研究という標本データの提供という課題に対しての研究がなされた。具体的にいえば、ミクロデータに関する統計的技法や社会制度上の問題などである。その後は、研究者から個票データ利用継続に対して強い要望があり「ミクロ統計データ活用研究会」につながったと触れている。

このように、標本データ提供に関する専門的な検証を終え、個票データ（匿名データ）の試行的提供へとつながっていく。厳格に秘匿処理をした 4 調査である、就業構造基本調査、社会生活基本調査、全国消費実態調査、住宅・土地統計調査の試行的提供が行われ、提供システムの形を模索した結果、これまでの目的外使用よりもこの試行的提供は利用しやすかっただけでなく、センターの存在がデータ利用者支援につながり、申請自体が円滑に進んだと述べている。同研究では、詳細は省くがこの試行的提供の実績を細かく分析し、その後 2007 年全部改正の統計法に際して整理をしている。その概要から、匿名データ提供の枠組みまで端的に述べ、最後に課題を付しており、匿名化をどの程度行うのが適切なのか、オンサイトによる利用の制約、事業所・企業調査の個票データの利用方法など 9 つを羅列している。それらを概観して言えることは、2019 年時点に考えられる課題と大方似通っているのではないかということであり、改めて公的統計の匿名データの利用体制を整える困難さがかがえる。

次に、2010 年代の研究では、星野 (2010) がミクロデータ提供制度の課題に焦点を当て、解決策を探っている。その課題に、利用目的制限の緩和と利用可能な統計調査の拡大の 2 つを挙げる。前者は、利用目的を問わず一般目的汎用ファイルの提供を提案している。後者に対しては匿名化が困難であるとされる事業所・企業調査も匿名化し、提供すべきだとしたうえで、具体的な事例をもとに模造の概念を提唱している。小林 (2012) は、2009 年（平成 21 年 4 月）に全面施行された新統計法にあった匿名データの提供を行う一橋大学の取り組みに着目している。研究当時、一橋大学では匿名データの提供とオンサイト利用の 2 つを行っており、その現状と問題点、展望を指摘している。とりわけオンサイト施設でのミクロデータ活用を検証したものとして、佐藤 (2017) がある。リモートアクセス型オンサイトで場所と時間の制約や成果物の持ち出しに審査があることを考慮しつつも、施設内で自由な探索的分析が可能となっていることを好意的に受け止めている。多くの研究が課題として示したうちのひとつである“匿名化”について研究をしたものとして伊藤 (2018) は、国際的動向を踏まえ分析している。具体的にいえば、フランス、アメリカなど欧米諸国の事例を参考に、法制度の側面と統計技術的側面から論じている。同研究は基本的に学術ユーザーによる二次的利用制度を論じたものであり、加工統計や政府内部の機関ユーザーによるプロジェクトを網羅した研究ではないが、主要国では日本よりも二次的利用制度の利便性が高いことを具体的に詳述している。

4. 法改正後ミクロデータの二次的利用制度を活用した学術研究の進展

統計法改正後、変化した点を考える上で、海外大学に所属している研究者によるミクロデータの利用が広がったかという点が挙げられる。これまで日本の調査票情報を利用した国際研究は、大変数が少ない。二次的利用制度全体でも留学生の研究、海外大学に所属する研究者（日本人も含む）と国内研究者の共同研究が毎年数件あるだけで、事実上二次的利用制度は海外在住の研究者には閉ざされてきた。これは二次的利用制度の利用条件に国内の研

究者という制約があり、海外研究者の場合には来日して日本国内で研究を終える必要があったためである。こうした研究に携われる外国籍の研究者は、ほぼ留学生や日本の大学に所属する研究者といった、日本滞在者である。

ただ、2019年5月の新統計法施行によってサテライトのサービスを利用できるなどマイクロデータ利用のハードルが下がったため、日本の公的統計データについて国内研究者と海外研究者の連携が活発化することが期待されている。ただし、統計法改正後、統計センターの公表実績によると、同センターが把握している海外の大学に籍を置いている研究者と日本国内の研究者の合同による調査票情報を用いた国際共同研究は、2019年5~12月までで3件であった。海外大学の研究者だけの申請事例は1件もないことから、法改正後も状況に大きな変化が無かったことが分かった。海外の大学研究者の場合は、日本につてが無いとこうしたデータにアクセスすることが難しい。希少な例外としては、日本が個票データの提供を行っている、ルクセンブルク所得統計やOECDで実施されている税務情報の交換といった非公表の成果がある。以降のサーベイに海外研究者による日本の調査票情報の利用事例は含まれていない。多国間での実証分析に日本のマイクロデータが利用される時代が来るにはまだまだ時間がかかると予想される。

上記の通り、マイクロデータの二次的利用制度（匿名データ提供、オーダーメード集計、オンライン利用）は新しい試みであり、したがって同制度を活用した学術研究も未だ試行錯誤を繰り返しつつ、徐々に利便性が高まっている。事例収集の一環として部会から西林が2019年11月15日に統計数理研究所にて開催された共同研究集会「官民オープンデータ利活用の動向及び人材育成の取組」に参加した。マイクロデータ利活用の最前線を知る上で重要な報告が数多く示されていた。以下は同研究集会をまとめたものである。

官民オープンデータ利活用の動向及び人材育成の取組（2019年度）の開催

日時：令和元（2019）年11月15日（金）9：45～17：50

会場：統計数理研究所 3階セミナー室1（D305号室）（東京都立川市緑町10-3）

https://www.nstac.go.jp/services/setumeikai_20191115.html

プログラム及び報告資料は上記URLにてすべて公開されている。この研究集会では、マイクロデータの二次的利用に関する学術研究に従事する研究者等によって、計19本の研究報告が行われた。特にその中では、匿名データ提供を利用した研究報告が目立った。

まず、匿名データ提供を利用した研究については、科研費研究の成果を発表した藤原翔（東京大学社会科学研究所）による「公的統計データを用いた社会学的研究の可能性とその教育」が挙げられる。この報告では、社会学において主要な関心の一つとなっている格差・不平等の再生産の問題を取り上げ、「公的統計は社会学者の関心に限定的にしか答えられない」と言われる一般的風潮の中、社会調査だけではなく、社会学研究として公的統計を活かしていく必要性（例えば、就業構造基本調査の職業、学歴情報から親の収入の推定など）が指摘された。また、澤野孝一郎（名古屋市立大学大学院経済学研究科）による『全国消費実態調査』のマイクロデータ分析—匿名データの利用による子どものいる世帯の家計の集計—では、「全国消費実態調査」の匿名データを利用している。同報告では、匿名データから子ども（15歳未満）がいる家庭の消費、エンゲル係数を全世帯と比較した結果、顕著な

差が見られなかったことが示された後、子どものいる世帯のみ詳細な分析を行っている。高橋行雄（BioStat 研究所）による「全国消費実態調査の匿名データ 4 年分の統合メタデータ作成」では、メタデータを参照し、4 年分をつなげて疑似ミクロデータを作成する統合方法が示された。また同報告では、今後の匿名データ利用活性化のために、データだけでなく、データ処理（共通の変数の作成）の方法論もセットで利用者に提供する必要性が指摘された。さらに、周防節雄・安井浩子（統計情報研究開発センター）による「国勢調査の続柄情報に婚姻状況・性別・年齢を加味して世帯員構成を直感的に表現できる変数の開発」では、昭和 45 年以降使われている家族類型ではない、新しい世帯構造を表現する変数を続柄（12 区分）・婚姻状況・性別・年齢を統合するためのアルゴリズムを開発したことが報告された。例えば、国勢調査等の世帯主について、誰を世帯主と見なすのかという判断は各世帯の主観に依存する。その場合に、ユーザーが自身のニーズに合わせて変数を作成することで、より実態に沿った研究を行うことできる可能性が指摘された。

オンラインの具体的事例に関する報告もなされた。オンラインを利用すれば、他のデータとの突合・照合が可能だが、1 時間の利用料が 4400 円と高額なことなど、数多くの課題が指摘されているのが現状である。椿広計（統計数理研究所）による「オンライン拠点の活用について—提供者視点から利用者視点へ—」では、オンライン拠点を活用した具体的な感想がいくつか紹介された。例えば、申請と持ち出し審査は迅速な対応だったこと、CSV ファイルとしての提供が可能であること、仮想 PC のメモリー 2GB では非力であること（社会生活基本調査について Excel ファイルを操作中にシンクライアント端末がフリーズした、Word ファイルの立ち上がりなどで端末がフリーズしたといった例）などの感想や、市区町村や県のマクロデータはオンライン環境に常時置いてほしい、国民生活基礎調査など未公開のデータを活用したい、などの要望も寄せられている。なお、統計数理研究所のオンラインを活用した研究事例として、椿氏自身が進行中である「エビデンスに基づく自殺問題の総合対策の確立に向けて」にも言及された。

また、オーダーメイド集計を活用した研究も報告された。森口千晶（一橋大学）・阿部修人（一橋大学）・井深陽子（慶応義塾大学）・稲倉典子（大阪大学）による「日本における所得階層別の栄養摂取と栄養素価格指数の長期的推計」は、家計調査のオーダーメイド集計データ（世帯主年齢別、世帯所得別、食品品目別の支出額と購入量）と日本食品標準成分表示を使用し、健康格差ではなく栄養格差に着目し、その長期的動向を把握しようとした研究である。栄養格差は 1981 年から 1995 年までは縮小し、その後はほぼ一定で格差の拡大はみられなかったということが示された。

さらに、公的統計としてオリジナリティに優れた成果として、芦屋恒憲（兵庫県企画県民部統計課・ビジョン課）による「兵庫県統計普及・加工分析事業の概要と課題」が挙げられる。同報告では、ユーザーから寄せられるデータ加工分析リクエスト（利用する統計表がわからない、専門用語が多いなど）への対応として、利用者のニーズ、分析目的に沿ってパッケージ化、事例の公表共有を進めていることが示され、地域統計が活用されるために何が必要かを整理している。なお、データの加工は大学および県が共同で実施している。

以上は研究集会に参加して直接まとめた内容であるが、概要については先述の通りミクロデータ申請内容が統計センターから公表されているため、公表資料が参考になる。この研究集会の報告等から二つの課題を指摘したい。第一に同研究集会ではミクロデータの国際

研究、国際共同研究に関する事例は報告されなかった。法改正後も依然として海外研究については課題が残ったままである。第二に、公益を認定した経済センサスほか母集団名簿の利用についてである。母集団名簿の利用は公的機関に限られていたが、5月の新統計法施行によって、大学など民間団体でも公益が認められれば、国勢調査、経済センサスといった母集団情報を直接利用できるようになった。しかし、同研究集会では、このようなケースは報告されなかった。公益認定の調査票情報利用のケースはかなりハードルが高いものの、統計調査の実査を外注で行うケースが近年大変多いことから、おそらくこうした実査を受注した機関が公益として認められて利用するケースは今後徐々に出てくるのではないかと予想される。

マイクロデータの種類に応じて個人情報やデータの規制が多くあり、専門的ハードルも高いため、マイクロデータの研究推進のためには研究集会でノウハウを持ち寄りながらハードルを一つ一つ乗り越える必要がある。現在日本ではオンサイトに代えてリモートアクセスを利用するため、試験運用に入っている。2020年1月28日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 (<https://kokucheese.com/event/index/589299/>) が開催され、研究成果が示される見通しである。先ほどの図表2、図表3のカナダのようにリモートアクセスに加えて行政データの匿名個票を利用する状況が日本に来るのはかなり先のことと思われるが、制約を乗り越えるよう懸命な努力が続けられている。

Ⅲ 公的統計と民間統計を組み合わせた研究成果

1. 消費に関するマクロ統計が示された背景

統計法改正の本来の目的は利便性を向上させ、公的統計と行政データをオンサイト機関で大規模に分析するといったことであった。実際にそうした状況にたどり着くのに時間がかかるとしても、萌芽的な研究は既にいくつか出てきている。本稿では省くが、東大の渡辺努研究室から生まれたナウキャストのようにPOSデータに基づく、日次物価指数の開発は毎日、日経新聞に載っているほか、しばしば経済財政白書で取り上げられるため、大変よく知られている。このような民間データから公益性を持ったデータを作り出す研究や、行政データと公的統計データを組み合わせるといった研究は今後日本でも数多く出てくると予想されている。その一つの成功事例ともいえるのは以下で取り上げる消費動向指数である。

ここでは消費に関する新たなマクロ指標を中心に、総務省などでの研究会の公開資料をもとにサーベイする。ビッグデータが報道番組でも特集が組まれるほどの話題になって久しい²⁾。ビッグデータとは、総務省によれば「事業に役立つ知見を導出するためのデータ」³⁾である。例えば、SNS (Social Networking Service) などのメディアへの書き込み、メール、GPSなどのセンサーデータや購入履歴などのウェブサイトデータなど、生活に密接に関係したものだと思われる。生活に関わっているからこそ、これまでの統計データとは違ったリアルさが反映されるともいえる。

2015年(平成27年)から経済財政諮問会議と統計委員会で、家計調査などの消費関連指標の改善が議論された。その後総務省での研究会は、ビッグデータなどを活用し、消費動向の全体構造を捉える速報性のある包括的な消費指標の体系を短期及び中長期視点に立って

開発し、さらにそれらを国民共有の統計指標として 2018 年（平成 30 年 1 月分）から提供することを提言した。

その後、2017 年（平成 29 年 7 月）には、「消費動向指数研究協議会」の設立を総務省が報道資料で発表した。構成員は、クレジットカード会社や小売、調査会社などの企業と研究者、機関から成る。ビッグデータの特性把握といった短期的計画から、新消費指標の定期公表という長期的目標を掲げ、公的統計の改善と高度化を目指すとしている。

補足として、実際の消費関連統計の現状として、一例を挙げる。まず、総務省が所管する家計調査が世帯単位のミクロ統計として行われている。他には、国全体のマクロ統計として民間の消費がおよそ半分を占める GDP 年次推計、速報などがある。GDP 統計は内閣府が所管している。

2. CTI マクロ、ミクロ

CTI マクロは主要素の消費部分、すなわち家計最終消費支出の月次情報等を推計する消費動向指数のことを指す。加工統計であるため速報性に欠ける GDP 統計に焦点を当てたものである。系列としては、名目・実質系列や財・サービス系列などが想定されている。

データソースには、バイアス補正したビッグデータを加える。調査データにそれらを融合させることで、推計モデルを構築し、高い速報性につなげていく。これまでと違い、データが得られた当該期だけでなく、推定値もよりの確に出すことができるとされる。

2019 年 9 月 6 日と 7 日に実施された経済統計学会では、民間企業のデータを使用した試算が報告された⁴⁾。そこでは、使用するデータに求められることやその利用にあたる留意事項、そして試算と検証が行われていた。第一に、使用するデータは項目ごとに十分なデータ量があることや入手のタイミングが公表に間に合う時期であること、かつ同質データの継続的入手可能性などが要件として提示された。次に、民間データを利用する上で、十分な把握が困難な項目があること、消費者の属性識別の困難さ、企業が保有する重複データによるダブルカウントの可能性や情報取り扱いへの配慮などが留意される事項として挙げられた。したがって、そういった多様な要件を満たすデータを入手していくことが今度の課題だと考えられる。最後に、POS データとクレジットカード情報を説明変数として利用した試算・検証に関して、結果の評価方法（要請）が 2 つ触れられていた。推定結果の安定性と GDP 統計との整合性である。前者は、推定結果が過去に遡って値が改定されていく統計の性質上、その改定幅は小さいことが望ましい。後者は、一次速報の GDP 統計の最新値に近いかつ同じ向きに動いているということだった。各評価方法に関して、指標も具体的に示されていたがここでは省略する。

CTI ミクロはマクロとは異なり、世帯の家計消費の月次動向等を各種データによる家計調査の補完と補強により推計していくものである。こちらも系列は CTI マクロと同じもののほかに、世帯類型別などが例示されている。具体的には、家計調査の改善と単身世帯把握のための単身モニター調査を行い、家計調査の上位モデルとなる指標体系を作成することを目指す（より詳細な流れは参考文献に挙げた資料に掲載されている）。なお、この指標は CTI マクロの説明変数としても用いられる。

3. 消費動向・データの現状・推定方法

推定方法の詳しいことは参考文献に譲ることにして、CTI マクロ、CTI ミクロのほかに、企業消費やインバウンド消費などの動向をビッグデータから推計することも考えられている。企業データとしては、クレジットカードやPOS データだけではなく、家計簿アプリや調査会社のもつデータを用いた指標構築を模索することになっている。

「状態空間モデル」に基づく時系列回帰モデルにより、GDP 統計の家計最終消費支出を被説明変数として、家計最終消費支出の月次動向を予測推計する。第一段階として、公的統計からトレンド、サイクル、ダミー項を抽出し説明変数を作成し、第二段階で家計最終消費支出の月次動向を推計するという流れになっている。

比較的新しい動きであるため、特化した先行研究自体は少ない。高部（2018）は、消費動向指数（とりわけ CTI マクロ）の推定モデルについて検討した。先に触れたようなマクロ消費動向指数に求められる点（月次変動を短期間で推測、公式の GDP 統計との整合性など）を満たすために、「状態空間モデル」の適用の利点を挙げて対応の可能性について言及した。その後、状態空間モデルの概要をまとめ、最終的に実際に開発途中のモデルを用い推定するという流れになっている。結果としては、マクロ消費動向指標の推定値は、月次の変動を捉えつつ、GDP 家計最終消費支出四半期支出の結果とも整合的であったとし、ある程度予測可能であることを示唆した。詳細なプロセスは当該論文に記載されている。

IV 新統計法以後成果が期待される分野～結論に代えて

2019 年 4 月までの旧統計法において母集団名簿の利用は公的機関に限られていたが、5 月の新統計法から大学など民間団体でも公益があると認められれば、国勢調査、経済センサスといった母集団情報を直接オンサイト施設にて利用できるようになった。ただし、オンサイトは 1 時間 4400 円の利用料がかかるなど、費用負担が大きいため、申請者はできるだけ政府自治体に研究の音頭をとってもらいインセンティブが働くかもしれない。政府自治体が調査票情報を利用する場合はオンサイトではなく、直接省庁に申請してデータを受け取る形式のため、法改正以前と変わらない。つまり、政府自治体は無料でサービスを受けられるのである。それでももし行政データやまとまった民間データが存在し、利用可能となるならば、オンサイトにデータを持ち込んで成果を出そうとする動きが出てくる可能性がある。先述の通り実際にアメリカはこの方式でセンサス局が税務データを使用しているとのことなので、自治体のデータをまとめて公的データと突合して行政サービスの効率を高めたり、アプリや IT サービスにつなげて利便性を向上させようとするなど、今後これまで日本で見られなかったようなダイナミックな研究が行われるかもしれない。

日本は人口も多く、公的統計の利用に多くの規制をしなければならないことからデータの制約に関する重要なブレイクスルーは公的統計分野よりも、同様に公益性のある民間データの分析からなされることが多い。例えば、渡辺（2016b）のナウキャストの事例が日本では大変よく知られているが、RESAS の成功やオープンデータの公開とアプリの開発といった内容は日本でも研究が報告されるようになってきた。

しかし、公的統計分野での研究は新統計法と運用の改善によって大きく進捗する可能性

を秘めている。オンサイト施設の利用はその端緒となる可能性がある。2019年5月30日に行われた、法政大学統計研究所による第18回国際ワークショップによると、アメリカ経済センサスでは、非回答の調査対象に対してデータ補完がなされるが、法人税データに基づいて推計されるとのことであった。櫻本が参加し、プレゼンを拝聴した他、アメリカセンサス局職員2名にも直接追加で質問し、情報を得ることができた。センサス局のブランディ・ヤーボロー氏によると、行政データと公的データの突合はオンサイト施設を通じて行う規則になっているとのこと、税務データの分析はこのオンサイトサービスで実現可能なことしか、できないという説明であった。一方、欧州やカナダでは匿名データの利用で既にリモートアクセスが広く使われるようになってきていて、先述のカナダのケースでは、実際に画面を開いてリモートアクセスに基づく匿名データの利用を目の前で見ることができた。匿名データファイルをダブルクリックして、申請済みデータを開くまで3秒ほどしかかからず、報告書を提出して利用を終えるのはその日中に終わるというスピードであった。しかも、加工統計の場合は事務系のシステム全般が利用可能な環境で、R、SASといった統計ソフトや応用的なソフトを自由に利用できるため、二次的利用制度は数時間の利用が一般的ということで、こうした点は現状数か月単位で手間がかかる日本とは雲泥の差である。カナダの場合、統計局内であれば、申請時に大義名分が無くても事後報告すら認められているので、所得税、法人税、GSTといった重要な税務データや政府の収支に関する行政データなど、政府内のデータ分析の利便性が事実上最大化されている。しかもカナダ-アメリカ間は北米自由貿易協定（NAFTA）に基づき、通関データを交換していたり、国境を越えたデータの交換すらできる仕組みを整えている。アメリカセンサス局のキンバリー・ムーア氏によると、興味深いことにNAFTAは公的統計に何の関係もないため、北米3国間の公的統計基盤の整備は法律とは関係のない自主的な取り組みだということだった。またEUも多国籍企業の個票情報を域内で流通させるデータベースを整えているし、オセアニアでもオーストラリア-ニュージーランド間での公的統計データの相互利用が行われている。

おそらく日本も統計法の厳しい制約を試行錯誤しつつ徐々に切り開き、こうした各国の状況に近づいていく可能性が高い。公的統計分野で利便性が高いデータのやり取りは、非公開であるケースも少なくない。GAFAと呼ばれる大規模IT企業の脱税問題に端を発してデジタル経済を税務で捕捉するBEPSプロジェクトのように、国際機関が非公開を前提に研究を実施している事例も多く、民間に情報が行き渡らないこともありうる。つまり、中国のように機関ユーザー向けに利便性が早めに整う一方で、大学などの一般ユーザーの利用に厳しい制限が付き続けることはありうる。その場合でも、一旦公的機関ユーザーの利便性が上がるのであれば、社会問題の解決のために民間ユーザー向けに徐々に規制が緩和されるきっかけとなりうる。

先述の通り、現状の日本には行政データのミクロデータを集中的に使用する機関や二次的利用を組織的に常時使用する機関ユーザーは統計組織に存在していない。しかし、このような専門性の高い機関が日本に設置され、利便性が高まれば、主要国のように行政データのミクロデータと公的統計のミクロデータを大規模に組み合わせるダイナミックな研究が生まれることが期待できよう。こうした努力は近年話題に上ることが多い、不平等の格差に関する研究など多くの社会的に重要な課題改善の糸口となる。例えば欧州ではトマ・ピクティ、アメリカではその弟子カブリエル・ズックマンがビッグデータや税務データから優れた研

究を次々と生み出している。新統計法施行をその端緒として、統計改革を進め、日本もこうした成果を生み出す基盤を整備できるか、今後の運用と規制の緩和が問われている。

注

- 1) なお、Ⅱの1～3、ⅢのCTIに関する内容は政府統計部会事業としてリサーチアシスタントの石田航一が調べ、部会メンバーが分担して取りまとめた内容が含まれる。
- 2) 鈴木良介（出演）、2012、「社会を変える“ビッグデータ”革命」、『クローズアップ現代』、NHK総合、2012年5月28日。などがある。
- 3) 総務省、2012、『情報通信白書』。
- 4) 総務省統計局、櫻井智章・高井健 / (独) 統計センター、深尾豊史・武藤杏里。

参考文献

・マイクロデータ利用に関して一覧

- 阿部穂日、2019、「公的統計マイクロデータの利用制度—改正について」経済統計学会東北・関東支部7月例会報告資料(立教大学池袋キャンパス)。
- 石田晃、1999、「アメリカ、カナダにおけるマイクロデータの現状について」、『研究所報』、No. 25, 1-34。
- 、1999、「オーストラリア、ニュージーランドにおける統計マイクロデータ提供の現状」、『研究所報』、No. 25, 49-68。
- 井出満、2004、「日本におけるマイクロデータ提供の現状」、『研究所報』、No. 32, 39-42。
- 伊藤伸介、2011、「わが国におけるマイクロデータの新たな展開可能性について—イギリスにおける地域分析用マイクロデータを例に一」、明海大学『経済学論集』、Vol. 23, No. 3, 36-54。
- 、2014、「イギリスにおける政府統計データの二次的利用の現状」、『ESTRELA』、No. 241, 10-20。
- 、2018、「公的統計マイクロデータの利活用における匿名化措置のあり方について」、『日本統計学会誌』、Vol. 47, No. 2, 77-101。
- 財団法人統計研究会、2018、「供給・使用表の枠組みの更なる活用及び四半期 GDP 速報の拡充に関する海外先行事例調査報告書（文献等の調査と訪問調査の結果）」内閣府 HP https://www.esri.cao.go.jp/jp/sna/seibi/kouhou/contents/pdf/180117_houkokusho.pdf
- 小林良行、2012、「公的統計マイクロデータ提供の現状と展望 一橋大学での取り組みをもとに」、『日本統計学会誌』、Vol. 41, No. 2, 401-420。
- 坂田大輔・鈴木雄大・櫻本健、2016、「海外公的統計のマイクロデータ利用に向けた取り組みの現状：国際マイクロ統計データベースの取り組みを中心に」、『社会と統計』、Vol. 2, 23-29。
- 佐藤朋彦、2017、「リモートアクセスを活用したオンサイト施設でのマイクロデータの新たな利用方法について」、『理論と方法』、Vol. 32, No. 1, 140-141。
- 渋谷政昭、1999、「マイクロデータの公有化と利用の技術的課題」、『研究所報』、No. 25, 101-

113.

総務省, 2019, 「ミクロデータの利活用も新たな時代へ」, 内部資料, 統計法改正前事前周知資料.

統計センター, 「統計センターが提供する公的統計のミクロデータ利用サービスについて」
統計センターHP 公表資料, 2020 年 1 月 6 日アクセス.

星野伸明, 2010, 「公的統計ミクロデータ提供制度の課題」, 『日本統計学会誌』, Vol. 40, No. 1, 23-45.

森博美, 1999, 「イギリスにおけるセンサスミクロデータの提供—マンチェスター大学による提供を事例として」, 『研究所報』, No. 25, 35-48.

———, 2004a, 「ミクロデータの利用特性と統計利用論」, 『研究所報』, No. 32, 9-38.

———, 2004b, 「政府統計ミクロデータの提供とわが国統計制度の今日的課題」, 『経済志林』, Vol. 72, No. 1・2, 33-65.

山口幸三, 2008, 「政府統計の個票利用と統計法改正—試行的提供の経験を踏まえて—」, 『経済研究』, Vol. 59, No. 2, 139-152.

・公的統計と民間統計を組み合わせた研究成果, ビッグデータ関連の文献一覧

岡崎陽介・敦賀智裕, 2015, 「ビッグデータを用いた経済・物価分析について」, 日本銀行調査論文.

(株)三菱総合研究所, 2018, 「ビッグデータを活用した行政課題解決に関する調査研究」,

https://www.stat.go.jp/dstart/research/pdf/2018_report_1.pdf (参照 2019-10-15).

総務省, 「統計法について」,

http://www.soumu.go.jp/toukei_toukatsu/index/seido/1-1n.htm (参照 2019-11-12).

総務省統計局, 2017, 「「速報性のある包括的な消費関連指標の在り方に関する研究会」報告(概要)」, https://www.stat.go.jp/info/kenkyu/sss/pdf/report_g.pdf (参照 2019-10-15).

総務省統計局, 2017, 「消費動向指数 (CTI) の開発に向けて」,

<https://www.stat.go.jp/info/kenkyu/sss/pdf/report.pdf> (参照 2019-10-15).

総務省統計局, 2017, 報道資料「「消費動向指数研究協議会」の設立」,

<https://www.stat.go.jp/info/kenkyu/ctikyogi/pdf/20170728.pdf> (参照 2019-10-15).

総務省統計局, 2017, 「消費動向指数 (CTI) について」,

<https://www.stat.go.jp/training/2kenkyu/pdf/gakkai/toukei/2017/sugita.pdf>
参照 2019-10-15).

総務省統計局, 2019, 「総消費動向指数 (CTI マクロ) の推定方法」,

https://www.stat.go.jp/data/cti/pdf/macro_ref.pdf (参照 2019-10-15).

総務省統計局ホームページ, 「消費統計研究会」,

<http://www.stat.go.jp/info/kenkyu/skenkyu/index.html> (参照 2019-10-15).

高部勲, 2018, 「消費動向指数 (CTI) : マクロ消費動向の推定について」, 『統計研究彙報』, No. 75, 21-40.

東大日次物価指数プロジェクトホームページ, https://www.cmdlab.co.jp/price_u-tokyo/ (参照 2019-10-22).

内閣府, 2018, 『平成 30 年版 経済財政白書―「白書」: 今, Society5.0 の経済へー』.

中村康治・河田皓史・田中雅樹・植前理紗, 2016, 「消費活動指数について」, 日本銀行調査論文, https://www.boj.or.jp/research/brp/ron_2016/data/ron160502a.pdf (参照 2019-10-15).

中村康治・三浦弘・丸山聡崇, 2016, 「消費活動指数の公表内容の拡充と見直しについて」, 日本銀行調査論文, https://www.boj.or.jp/research/brp/ron_2016/data/ron161007a.pdf (参照 2019-10-15).

藤田隼平, 2017, 「POS データを用いた経済分析の試み―小売価格と景気動向との関係性の検証」, 『経済財政分析ディスカッション・ペーパー』, DP/17-4.

丸山歩・嶋北俊一・落合牧子・上田聖, 2015, 「CPI と東大物価指数の乖離の分析について」, 『統計研究所彙報』, Vol. 72, No. 4, 55-78.

渡辺広太・渡辺努, 2013, 「スキャナーデータを用いた日次物価指数の計測」, 『東京大金融教育研究センターワーキングペーパー』, CARF-J-094.

渡辺努, 2016a, 「店舗別インフレ率から読み取れること」, ナウキャスト『マンスリーレポート 2016 年 2 月号』.

渡辺努, 2016b, 「東大日次物価指数プロジェクト」, 総務省統計局資料, https://www.stat.go.jp/info/kenkyu/sss/pdf/161014_shiryoku1.pdf (参照 2019-10-22).

・その他

「毎月勤労統計の調査ミス, 厚労省, 把握後も公表せず」, 『日本経済新聞』, 2019 年 1 月 9 日, 朝刊, p.5.

「労働保険を過少支給か, 毎月勤労統計の調査ミス, 影響広がる」, 『日本経済新聞』, 2019 年 1 月 10 日, 朝刊, p.5.

「「基幹統計」信頼揺らぐ, 勤労統計不適切調査, 04 年から放置」, 『日本経済新聞』, 2019 年 1 月 11 日, 朝刊, p.5.

「統計委, 不正再発防止策を建議 調査法公開やデータ保存」, 『日本経済新聞』, 2019 年 9 月 30 日, 電子版 (参照 2019-12-03).

「統計不正防止へ対策, 総務省に支援窓口, 政府素案, 調査票を一元保管」, 『日本経済新聞』, 2019 年 11 月 22 日, 朝刊, p.5.

「統計職員に専門資格を創設 政府, 不正の再発防止で」, 『日本経済新聞』, 2019 年 11 月 20 日, 電子版 (参照 2019-12-03).

Summary

The Recent Development of Micro Data Analysis and Official Statistic Databases in Japan

Takeshi Sakuramoto
Shogo Nishibayashi
Shinichi Hamamoto

The Statistics Act which was completely revised in 2007 was partially revised again and put into force in May 2019. Generally speaking, law systems are not changed all at once, but they are gradually developed through feedback processes between research analysis and law system changes. Because statistic systems in Japan are far less accessible compared to some Nordic countries, it is very important for the Japanese system to catch up with those countries through law revision and improving the execution of laws.

In this paper, we will show survey research about statistical studies using micro data following the Statistics Act revision in 2019. In addition, we will use not only literature review but also the reports from our research trip. It was found that there are only a few studies which advance official statistic systems focusing on on-site data usage services, utilize established frame databases such as the Economic Census, and contain a combination of other official statistics and private data. Therefore, it can be determined that the accessibility of statistic systems in Japan was not improved by the law revision in 2019 and the gap between Japanese systems and Nordic systems continue to exist for the moment.

It was said that on-site data usage services would be gradually expanded. Considering the secondary use services for official statistics have been well developed, if administrative data availability is improved, it would be possible that the recent development of micro data analysis would contribute to the advancement of statistic systems.

Key words : secondary use services for official statistics, Statistics Act revision, micro data, on-site data usage service, remote access data usage service

(44) 社会情報教育研究センター研究紀要『社会と統計』第 6 号

《論文》

データサイエンス教育の現状と課題 —高等教育機関におけるデータサイエンス教育の方向性—

山口 和範
山口 誠一
門田 実

【要旨】 近年、高等教育機関におけるデータサイエンス教育の議論が盛んである。統計学部や学科が存在した海外はもとより、統計という名の付く学部学科がなかった国内においても、データサイエンス学部が複数の大学で開設された。さらに、国や産業界からは、数理・データサイエンス教育の必修化という要望も出されている。一部の大学では、全学生を対象に、データサイエンス関連の科目の必修化を進めている。一方、データサイエンスという領域やデータサイエンティストという呼称については、様々な定義や論点があることも事実である。立教大学では、2018年度から、これまでの社会調査や統計を基礎としたデータサイエンス副専攻を、全学生を対象にスタートさせている。本論文では、これまでの統計教育改革の流れを踏まえた上で、データサイエンス教育の現状と今後の課題について議論する。なお、ここで議論するデータサイエンス教育のターゲットは、リテラシーレベルの教育としての学生とする。

キーワード：統計教育、統計的思考力、コミュニケーション、課題解決

I はじめに

1990年代に入り、初等中等教育から高等教育に至るまで、海外での統計教育再編の動きが活発化した。その背景には、国家的に推進される科学技術振興政策があり、これが諸外国の学校教育の中で、統計教育の方法論として新しい枠組みであった“Statistical Thinking（統計的思考力）の育成”という柱をもたらした。先進各国は、人材や技術など「知の創造」をめぐる大競争時代に突入し、世界全体での持続的発展や自国の産業競争力の国際的優位性の獲得を目指し、科学技術・学術研究の戦略的な推進政策を推し進めた。とくに、1998年、全米研究会議が通称オドム・レポートを取りまとめ、数学と他分野および産業との連携の重要性を指摘して以降、米国科学財団（NSF）は、重点領域に数理科学を採用し、その中の重要テーマとして、“巨大データに関する数学的・統計的挑戦”、“不確実性の管理とモデリング”、“複雑な非線形システムのモデリング”を挙げた。このような振興政策を支える人材の育成のため、NSFが重点領域に数理科学を採用した時点でその具体的な目標として、

- ・数学および統計学の基盤研究の推進
- ・他の自然科学・工学との協働研究の推進
- ・学校教育における数理科学教育の推進

を掲げており、3点目の数理科学教育については、学校教育における主に数学や理科の中に反映され、前者2項との連動性から、数学や理科教育の中で統計教育の比重が増している。さらに、その中の統計教育においても単純に統計リテラシーを有しているだけでなく、数学以外の他の学問領域、また産業や市民生活の場における科学化（諸対象を科学的に探求するプロセス）を促進できる人材、いわゆる統計的課題解決型の思考力を持った研究者や生活者の育成をも強く意図されている。統計的思考力が科学技術発展の第3の腕と位置付けられ

る中で、統計科学の研究領域の中においても、統計的思考力自身の定義やその育成方法の研究が活発化した。大学における統計学入門教育の内容に関しては、1996年に米国統計学会（ASA）と全米数学協議会（MAA）の共同カリキュラム委員会がデータ分析の実践の要素を盛り込んだ統計教育の共同指針を発表し、翌1997年には the College Board が Advanced Placement テストに統計科目をこの指針の下に導入し、高校生の統計 AP テスト受験者数が毎年増加をたどり、10万人を超えるまでに至っている。この学校教育における統計教育の需要により、米国統計学会は、新しい枠組みでの統計教育のガイドラインと評価方法を初等中等教育から大学の統計入門コースまで体系的に作成・公開し、現場の教育を積極的に支援している。中でも、GAISE（Guidelines for Assessment and Instruction in Statistics Education）レポートでは、学校教育レベルと高等教育レベル、それぞれにおいて、教育の内容および評価についての指針を提示し、その後の統計教育の改革に大きな役割を果たしている。

一方英国では、王立統計学会（Royal Statistical Society）が統計教育センター（RSSCSE: Royal Statistical Society Centre for Statistical Education）を設立し、英国の統計局との協力体制の下で、具体的な教材開発、大学における統計教育内容の指導と認証など統計教育のサポートを組織的に行っている。他の諸外国も同様の統計教育改革が進んでいる。

日本においても、日本統計学会が中心となり統計検定がスタートし、その受験者数も年々増加するなど、統計の学びに関する需要も高い。また、ビッグデータをキーワードとしたデータサイエンティストへの産業界からの関心もあり、高等教育機関における統計思考力養成への期待は高い。渡辺（2013）は、日本における統計教育改革の方向性を、問題解決のための重要なツールという視点で議論している。また、山口（2015）では、問題解決のために必要なリーダーシップ教育との関連で、統計教育の新たな枠組みについて議論を行い、その具体例として、立教大学における取組事例を紹介している。

このような統計教育改革の流れにさらに拍車をかけたのが、人工知能の社会実装の具体化の進展と、データを无尽蔵に生み出される有益な資源ととらえ、それを有効利用できる人材の必要性への気づきである。そこでは、統計という枠組みのみで捉えるのではなく、情報技術の領域等も広く含んだデータサイエンスという枠組みでの教育改革が求められている。本論文では、現状でのデータサイエンス教育の方向性をまとめ、今後の在り方について議論を行う。

II データサイエンス教育

1. データサイエンス教育の進展

データサイエンスというキーワードは、そのキーワードを用いる専門家の専門領域により重きを置く比重は変わるにせよ、統計学、情報技術、および、適用する専門領域の知識を活用して、問題解決にあたる方法論であると考えるのが一般的であろう。

データサイエンスは、科学や産業に革命を起こす新しいアプローチとして期待されている。これは、多くの研究分野のみならず社会の至るところで大規模で網羅的なデータが急速に蓄積されるようになり、データやその分析結果を利用することが可能になることと、利用

しないことが一種の機会損失であるという意識の高まりによると思われる。このことは、今後データを経済、社会及び日常生活の発展と新しい価値創造の源泉となるように有効利用できる人材が必要不可欠であると考えられている。

日本においては、2015年に情報・システム研究機構ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会の報告書『ビッグデータの利活用のための専門人材育成について』(北川 2015)で、「データから新しい価値を生み出すことができる「棟梁レベル」をはじめとするデータサイエンスの専門人材を育成することの重要性は言うまでもない。同時に、専門人材が生み出す成果の受け手となる国民全体のデータリテラシーを醸成することが、我が国においてデータサイエンス、AIを活かして超スマート社会を実現していくために必要不可欠だと考えられる。したがって棟梁レベルからリテラシーレベルまで裾野の広い人材育成が望まれる」と述べられている。

リテラシーレベルでの教育をここでの議論の対象とすると、データリテラシーとしては、データサイエンスやAIの不適切な利用や、個人データの利用から生じうる倫理的問題を認識できるようになることの重要性も広く言論されている。これは、その認識が今後のデータ駆動型社会を安全・安心に過ごすために必要不可欠なことであるからである。もちろんデータサイエンスの専門人材も、そのような側面を十分に認識した上で価値を創造することが強く期待される。

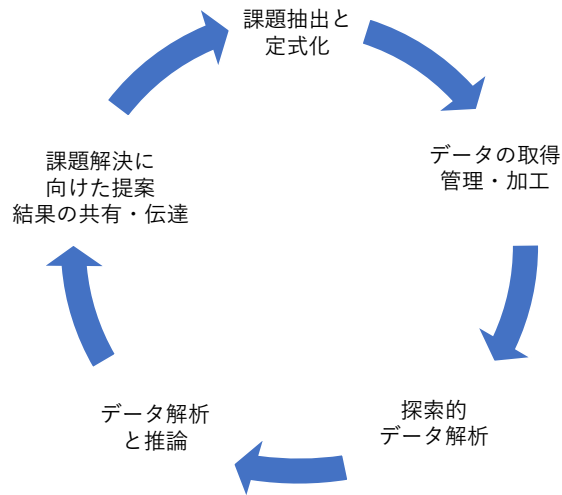
このような議論においては、統計教育における市民が身につけるべき内容の議論を行った Utts (2003) が十分に参考になる。ここでは、確率統計を学んだ市民が身につけておくべき内容として、「無作為化実験と観察研究の違い：交絡の理解」、「統計的有意であることの意味：大標本の場合」、「有意でないことの意味：小標本の場合」、「バイアスが生じる原因」、「偶然の意味：めったにないこともどこかで起こる」、「因果の方向性と条件付き確率」、「バラツキの理解」の7点を挙げている。

データサイエンスにおけるデータの状況は、20世紀に発展した統計的推測法とは、一部その規模や形式において異なる点はあるものの、ここでの指摘の内容については、データサイエンスという呼称においても、少なくとも身につけておかないといけない統計学的側面のスキルであると考えられる。

一方で、データサイエンスは単なる統計的思考力を身につけることというより、それを課題解決に向けて組織やチームとしてデータに基づく知識発展のサイクルを回すことを強く意識する点が重要である。De Veaux et al. (2017)で指摘されているように、データサイエンスにおいては

課題抽出と定式化
データの取得・管理・加工
探索的データ解析
データ解析と推論
結果の共有・伝達、課題解決に向けた提案

というプロセスがあり、課題解決と同時にデータからの知識発見や新たな課題発見のためのサイクルが重要である(図表1)。



図表 1：データサイエンスにおける課題解決のサイクル

ここで、図表 1 のようなデータサイエンスのサイクルを理解して実行する際、個人での実行というよりチームや組織での実行という意識が必要となることを強調しておきたい。

アメリカの The National Academies of Sciences, Engineering, and Medicine によるデータサイエンスの学部教育に関する報告書「Data Science for Undergraduates Opportunities and Options」では、学部教育において下記の 10 領域を重要なものとしてあげている。

- ・ Ethical problem solving 倫理に配慮した課題解決
- ・ Data description and visualization データの記述・可視化
- ・ Data management and curation データの取得・管理・加工
- ・ Statistical foundations 統計の基礎
- ・ Mathematical foundations 数学の基礎
- ・ Computational foundations コンピュータの基礎
- ・ Data modeling and assessment モデリングと評価
- ・ Domain-specific considerations 適用領域の知識
- ・ Communication and teamwork コミュニケーションとチームワーク
- ・ Workflow and reproducibility ワークフローと再現性

ここで、コミュニケーションとチームワークに注目したい。複雑化した課題に向き合う際に、1つの学問領域や成功者のこれまでの成功体験により、課題解決を図るのではなく、多様な人材からなる一チームでデータに基づく課題解決を目指す際に、求められるのがコミュニケーションとチームワークであり、言い換えればリーダーシップである。

このようなコミュニケーションとチームワーク、リーダーシップは、チームや組織での課題解決のためのスキルとして様々な領域で意識されてきた。

高等教育機関においてアカデミックな教育の重要性に加え、近年社会への人材輩出組織としての、人材育成としての教育の重要性が指摘されたのはそう古くはない。日本においては、キャリア教育の意識の高まりとともに、社会人基礎力や人間力というキーワードが大学教育の改善の方向性を示してきた。これは、学生が専門として学び身に付けた実力を、社会において実践する力ともいえよう。現在、高等教育機関において機能別分化が求められるなど、高等教育機関に求められる人材育成の内容が変化してきた。これまでのアカデミックな教育を行うという視点に加え、産業界等での働き手としてのプロフェッショナル教育の視点が重視されている。ひとつの工夫は、大学内に閉じた教育ではなく、社会と繋がりを持った学びの場の提供が求められており、産学連携による教育がある。これは、グローバル化への対応と共に、現代社会における高等教育機関につきつけられた大きな課題である。データサイエンス教育重視への流れもその一環といえなくもない。

一方、統計家育成においては、米国をはじめとする各国で **Professional Statistician** の資格制定が行われてきたが、その要件としてコミュニケーション力やリーダーシップスキルをあげる場合がある。統計の専門家だけで問題解決が図られるわけではなく、分野の専門家と協力して課題解決に当たる際の必要な能力として、コミュニケーション力やリーダーシップスキルが必要となる。リーダーシップスキル養成のためのプログラムは、ビジネススクールには設置されていることが多いが、米国の **University of North Carolina** の生物統計のコースでリーダーシッププログラムがスタートしている。さらに、**American Statistical Association (ASA)**でも、学会としてリーダーシップ教育の重要性を認識し、新たな取り組みをスタートさせている。

近年のデータサイエンス教育において、「ビッグデータ」のキーワードは重要な視点である。多種多様で巨大な規模のデータがスピードを伴い提供される環境で、その情報をどう生かし、問題解決にどう結び付けるかが大きな課題である。ここでの重要な視点は、これまでの統計教育に加えどのような内容を付加すべきかということであろう。Utts (2003) の指摘した市民が身につけるべき内容は、前述のとおり時代が変わっても、データに基づく意思決定を行う際には必ず必要となる事項である。とくに、確率についての常識や統計的有意性の正しい理解は、ビッグデータ時代こそ、その重要性が高まると思われる。

一方、椿 (2013) はマネジメントサイエンスにおける統計科学の役割を指摘するとともに、ビッグデータ時代に必要とされる専門家層の育成についての問題提起を行っている。統計学部や統計学科が存在しない日本の特殊事情を踏まえ、人材育成への対応が喫緊の課題という指摘である。佐々木 (2013) は、ビジネスにおけるデータ分析のプロフェッショナルの必要性と、そこに求められるビジネススキル、分析スキル、IT スキルのバランスの重要性を示すとともに、バランスのとれた人材の確保の困難性を指摘している。さらに、人材不足の具体的な解決策の方向性を示すとともに、立教大学経営学部での産学連携のデータサイエンティスト養成講座の試行の紹介も行っている。そこでは、統計の理論学習に重点を置くのではなく、実際のデータと分析のためのツールを使い、グループでのディスカッションを含むグループプロジェクトとしての実践を行っている。この点で重要なことは、個人での分析力の向上ということだけでなく、グループとして分析を行い意思決定につなげられるかという点であろう。米国の統計学会が認証しているプロフェッショナル・スタティスティシャン (P-STAT) の資格でも、統計分析の知識や技術力に加え、コミュニケーション力

やリーダーシップが求められている。ビッグデータ時代において、その傾向は、一段と高まると思われる。データの多様性や規模が拡大するほど、多様な専門性を持った専門家の協働の必要性が増大するであろう。チームとしての問題解決を図る上で、コミュニケーション力とリーダーシップは不可欠で、そこに分析力を組み合わせることが重要と考える。

2. データサイエンス教育における倫理教育

データサイエンス教育では、統計学に基づく分析手法を身につけるだけではなく、コンピュータで実データを処理するスキルやその際の注意点を習得することが重要である。また、ネットワークを介したグループでの共同作業により、データサイエンスの課題解決を導くことも想定され、ネットワークに繋がったコンピュータを扱うにあたり情報倫理の基礎を理解し、情報セキュリティへの理解も求められる。

情報倫理に関しては、知的財産に関する法律、個人情報保護に関する法律、不正アクセス行為の禁止等に関する法律などを含めた法規が情報社会で果している役割への理解、情報社会で生活する上でのマナー、モラル、倫理の意義や、法を遵守することの重要性の理解、さらに、現代情報社会では技術革新のスピードが非常に早く、法規や制度の対応が間に合わないケースがあること、そのために個人のモラルや倫理に基づく正しい対応が期待されることへの理解も含まれるであろう。

一方、データに関連する法律・規制も学ぶ必要がある。立教大学社会情報教育研究センターには、政府統計部会が設けられ、公的統計への理解を進める学びの場を提供しているが、個人情報保護法や統計法はデータに関する基本的な法規であり、後者では、公的統計が行政利用だけではなく、社会全体で利用される情報基盤として極めて重要であり、根拠に基づく（データに基づく）政策立案の基礎であることを学べるよう努めており、これもデータサイエンス教育における重要な要素である。

データサイエンスを応用し多種多様で大量のデータを分析した成果は、社会に多大なる便益をもたらす。その一方で、深刻な倫理的な問題を引き起こす可能性があることが指摘されている。例えば、個人情報の曝露や、プロファイリング等の自動処理による年齢・民族・性などで層別された特定のグループへの差別や人権侵害である。ここでは特に、倫理に配慮したデータ収集や利活用、匿名化されたデータであっても個人情報を曝露するリスクがあることへの理解を求める必要がある。

さらに、データの収集においても、倫理に配慮しなければならない。立教大学では社会調査士資格を取得するための科目設置が多く学部で行われており、その科目群のなかでは、調査における倫理教育が行われており、その意識は一定程度浸透していると思われる。ただ、人工知能の社会実装の場を中心として、データの活用における新たな倫理問題も台頭しており、その対応が大きな課題といえる。

3. 立教大学におけるデータサイエンス教育

立教大学では、2016年からグローバル教養副専攻がスタートし、社会情報教育研究センターが中心となり、2018年度からデータサイエンス副専攻がスタートした。このデータサイエンス副専攻の提案書では、「2013年6月に閣議決定された『日本再興戦略-JAPAN is BACK-』の「4. 世界最高水準のIT社会の実現、⑥産業競争力の源泉となるハイレベルな

IT 人材の育成・確保」で述べられている“IT やデータを活用して新たなイノベーションを生み出すことのできるハイレベルな IT 人材”の育成を目指す」とあり、「具体的には、データ活用力や IT 技術を身につけることで、メジャーとしての専門性をよりグローバルに活用できるための副専攻プログラムを目指す。新たなイノベーションの担い手としての専門性を重視しながら「データ活用力や IT 技術を身につける」副専攻を用意する」としている。このデータサイエンス副専攻では、指定された 16 単位の単位修得が求められる(図表 2)。

履修モデル: <テーマ>Data Science(データサイエンス)

系列	目的	科目	必要 単位数
第1系列	調査の仕組みを学ぶ	多彩な学び「社会調査入門」「社会調査の技法」 学部展開科目*1	4単位 以上
第2系列	データ活用の基礎的な力を身に付ける	(2-1)基礎科目 多彩な学び「データ分析入門」「データの科学」「数学の世界」 学部展開科目 *1	8単位 以上*2
第3系列	データ活用の発展的な力を身に付ける	(3-1)基礎科目 多彩な学び「Introduction to Statistics 1, 2」	
海外体験	海外を知る	(3-2)先端科目 2019年度開設科目はなし	認定
修了に必要な総単位数			16単位

*1 学部展開科目における履修条件(他学部履修の可否、先修規定等)は、科目設置学部の履修規定等に基づく。

*2 第2系列および第3系列の基礎科目(区分2-1、3-1)から4単位以上、第2系列および第3系列の先端科目(区分2-2、3-2)から4単位以上をそれぞれ修得すること。

1

図表 2 : 立教大学データサイエンス副専攻

展開されている科目としては、2019 年度で、第 1 系列が 10 科目、基礎科目は 61 科目、先端科目 38 科目となっている。なお、英語での開講科目は、基礎科目で 2 科目が展開されており、今後数科目追加される予定である。また、7 科目はオンデマンド科目として展開されている。もちろん、池袋と新座の両キャンパスそれぞれで、十分な科目が展開されており、副専攻修了はキャンパスを問わず可能となっている。

立教大学で実施しているデータサイエンス副専攻における具体的な科目構成であるが、社会調査や統計学、データ分析関連の科目群と、公的統計の利活用に関する科目などで構成されている。前節で述べた内容全般をカバーできる内容とはなっていない点やデータに基づく課題解決を含む実習系の科目がない点などの課題がある。しかし、グローバル教育センターが提供しているグローバルリーダーシッププログラム (GLP) との連携などが今後期待できる一方、倫理面の教育の強化は必須であると思われる。大規模データからの予測と個人情報との関連などをきちんと意識しながら、課題解決につなげられることが重要で、そのような内容の学びの場を、機械学習等の新たな分析手法の学びの場の提供と合わせて行う必要がある。

III まとめ

データサイエンス教育の重要性は、社会におけるデータの普及や活用の推進、さらにエビデンスに基づく意思決定が重視されるほど増すものと思われ、そのための人材育成が課題となる。日本においては、統計学部や統計学科がないことが問題として指摘されてきた。しかし、滋賀大学をはじめとするデータサイエンス学部の創設や、立教大学で実施しているデータサイエンス副専攻の実施は、今後大いに広まっていくと推察される。その中で、課題発見課題解決のための1つの方法論としてのデータサイエンスの位置づけはまだ定めてはいない。また、国が主導して議論が進んでいるデータサイエンスのモデルカリキュラムについての議論や実施もこれからである。大学におけるデータサイエンス教育は、今後広まっていくことに疑問の余地はないが、その教育の質保証や教員確保など課題も多い。特に、数理重視ではなく、課題解決につながる統計的思考力や一定のコンピュータスキル、さらにコミュニケーション力やリーダーシップスキルの育成と、必要となる倫理観の醸成が重要となるであろう。そのために、グループや組織として取り組む実践の場の提供が重要なポイントとなる。このことから、高等教育機関と産業界が協働してのデータサイエンス人材育成の取り組みが重要となるであろう。

参考文献

- 北川源四郎, 2015, 『ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会報告書』, 大学共同利用機関法人情報・システム研究機構.
- 佐々木宏, 2013, 「ビッグデータ・アナリティクスの組織適用とデータサイエンティスト」, 『経営システム』第3巻第4号, 237-241.
- 椿広計, 2013, 「ビッグデータ時代のアナリティクスーデータの価値を増大させるヒトとコト」『経営システム』第3巻第4号, 218-223.
- 山口和範, 2014, 「大学におけるデータサイエンス教育」『品質』44巻3号, 日本科学技術連盟.
- 山口和範, 大橋洸太郎, 大川内隆朗, 丹野清美, 2014, 「データ活用力の育成を意識した統計教育」『大学教育と情報』2014年度No.4, 32-35.
- 渡辺美智子, 2013, 「知識基盤社会における統計教育の新しい枠組み ～科学的探究・問題解決・意思決定に至る統計思考力～」『日本統計学会誌』第42巻第2号, 253-271.
- De Veaux, R. D. *et al*, 2017, Curriculum Guidelines for Undergraduate Programs in Data Science, *Annual Review of Statistics and Its Application*, 4, 15-30.
- Utts, J. 2003, What Educated Citizens Should Know about Statistics and Probability, *American Statistician*, 57(2): 74-79.

Summary

Challenges of Data Science Education : The Direction of Data Science Education in Higher Education

Kazunori Yamaguchi
Seiichi Yamaguchi
Minoru Kadota

In recent years, there has been much discussion of data science education in higher education. Data science departments have been established at several universities in Japan, where there were no departments with the name “statistics.” In addition, the government and various industries have requested that math and data science education be compulsory. Some universities are making data science courses required subjects for all students. Rikkyo University began a data science minor program based on social research and statistics in 2018. However, there still exists various definitions and issues regarding the area of data science and data-scientists. In this paper, we discuss the current status of data science education and the future issues based on the recent trend of statistical education reform. The topic of data science education discussed here is regarding students receiving literacy level education.

Key words: statistical education, statistical thinking, communication skill, problem solving

《研究ノート》

『日本帝国統計年鑑』を電子化しての雑感

五十嵐 彰

【要旨】 本稿では、1882 年（明治 15 年）から大戦によって中断される 1940 年（昭和 15 年）まで毎年継続的に収集された『日本帝国統計年鑑』について、データとしての特性と重要性、筆者が実際に電子化したデータの紹介、データを実際に電子化しての問題点、最後にその他の有用な戦前データと今後の展望について紹介する。今回は都道府県別外国人数をはじめとする 14 変数を電子化し、戦前の外国人数の推移について知見を得た。しかしながら、同時に電子化には多くの困難があり、それを 1) 形式、2) 欠損、3) 一貫性の 3 つの観点から整理する。

キーワード：日本帝国統計年鑑、戦前、外国人数

I 本稿及び『日本帝国統計年鑑』概要

本稿では、『日本帝国統計年鑑』のデータとしての特性と重要性、筆者が実際に電子化したデータの紹介、データを実際に電子化しての問題点、最後にその他の有用な戦前データと今後の展望について記述する。

『日本帝国統計年鑑』とは、1882 年（明治 15 年）から大戦によって中断される 1940 年（昭和 15 年）まで毎年継続して集められた統計情報をまとめた資料である。全 59 巻が刊行されており、図表 1 に示したような多岐にわたる統計情報を含んでいる。図表 1 では比較のために項目の順序を変えたが、多くの項目を第 1 巻から第 58 巻まで通して収集しているのがわかる。細目は省いたもののこれらの小項目も充実しており、例えば第 1 回の小項目を足し合わせると 279 項目、第 29 回では 604 項目、第 58 回では 416 項目となっている。小項目の中には例えば県別の牛車数（牛一頭引きか二頭引きかの区別あり）や難破船の数、病気ごとの患者数といったものまである。

戦前の日本に関する最も重要なデータといえるが、紙のデータとして保管されており、管見の限り電子化はなされていない。国会図書館デジタルアーカイブには画像情報として保存されたものがあるが、あくまで画像情報であり、かつ 30 巻以前のものは確認できない。電子化する必要のないデータであれば、現状でも特に問題ないといえる。しかしながら殊『日本帝国統計年鑑』に限っては、重要性が認識されているにも関わらず電子化が進んでいないといえる。例えば『日本帝国統計年鑑』は海外の経済史研究に活用されており、こうした研究のいくつかはトップジャーナルに掲載されている（e.g., Grossman & Imai, 2008; Tang, 2014）。重要な知見を引き出すことのできる可能性を秘めているにもかかわらず、十分に活用できるよう未だ整えられていないといえるだろう。

II 実際のデータ

筆者が実際に入力・分析したデータを実例として記す。入力の際には社会学研究科大学院

生の小松恵さんと宮澤篤史さんに大いに助けてもらった。この場を借りて改めて感謝申し上げます。

項目内容	第1回	第29回	第58回
土地	土地	土地	土地
天候		気象	
人口	人口	人口	人口
第一次産業	農業 山林 漁業及び製塩	農業 山林及び狩猟 漁業及び製塩	農林及び水産
第二次産業	鉱山 工業	鉱業 工業及び賃金 築造	鉱業及び工業 労働及び職業紹介
第三次産業	通運 銀行及び金融 外国貿易	陸運 水運 郵便及び電信電話 内国商業及び会社 銀行及び金融 貯金及び保険 外国貿易	交通 商業及び金融 貿易
病院・衛生	衛生	衛生	
宗教	社寺	社寺及び教会	教育，神社及び宗教
教育	教育	教育 教育および慈恵	社会事業
治安	警察 監獄 司法	警察 監獄 裁判及び登記	警察，衛生及び災害 司法
軍	陸軍 海軍	陸軍 海軍	
政治	政事	議員選挙 官吏公吏及び恩給 財政	選挙，官吏公吏及び恩賞 財政 産業国体
財閥等	財閥	爵位勲章及び褒章	
開拓地	北海道	北海道	
植民地・領土		朝鮮 台湾 樺太	

図表1：3時点における『日本帝国統計年鑑』掲載内容

今回は自分自身の専門である外国人研究のために『日本帝国統計年鑑』のデータを入力した。都道府県レベルの変数のみを用いている。研究目的としては、外国人の居住行動の規定要因を探ることであり、なぜある地域に外国人が多く居住するのかを検討することにある。外国人の数が増えることにより、経済や福祉など地域の状態が（良い方向にも悪い方向にも）変わると言われてきており、そのため外国人がどこに住むかを検討することは重要な課題であった。多くの先行研究では、事前の外国人住民の存在が新たな外国人移住者を生むことがわかっていて、しかしながらこれらの研究は、ではそれらの事前の外国人住民がどのように居住地選択をしたかという問いに答えることは出来ない。こうした問いに答えるには、外国人ネットワークがない（もしくはほぼない）状態にまで遡る必要がある。日本では鎖国解禁以後も外国人の居住地域を複数の地点に限定しており、1899年に至るまで外国人はこれらの地域以外に住むことは原則許されなかった。『日本帝国統計年鑑』には1899年およびそれ以降の県別外国人数が記載されており、この情報を使うことにより、外国人ネットワークが（ほぼ）存在しない状況において、なぜ外国人がある地域に移り住んだのかという問いに答えることができる。このようなモチベーションのもと、小松さんと宮澤さんにデータを入力してもらった。実際に集めたデータは図表2の通りとなる。

	年	欠損年
外国人数	1899-1938	1912, 1914, 1925
日本人人口	1898-1938	-
男子不就学率	1898-1937	-
女子不就学率	1898-1937	-
犯罪者検挙数	1898-1938	1924
米生産高	1898-1938	-
甘藷・馬鈴薯生産高	1898-1938	-
製糸工場数	1898-1937	1906
反物・帯値段	1898-1937	-
農業企業数	1898-1938	-
商業企業数	1898-1938	-
工業企業数	1898-1938	-
水陸運輸業企業数	1898-1938	-
所得税額	1902-1938	-

図表2：『日本帝国統計年鑑』収集データ一覧

実際に収集できた年と、データが完全に欠損となっている年を示した。この他にもいくつかの県が欠損となっている場合がある。外国人数については、制度として外国人が開港地・開市地以外に居住できるようになったのが1899年であり、その年からの収集となっている。それ以前のデータは開港地・開市地のみのものであり、今回は収集しなかった。

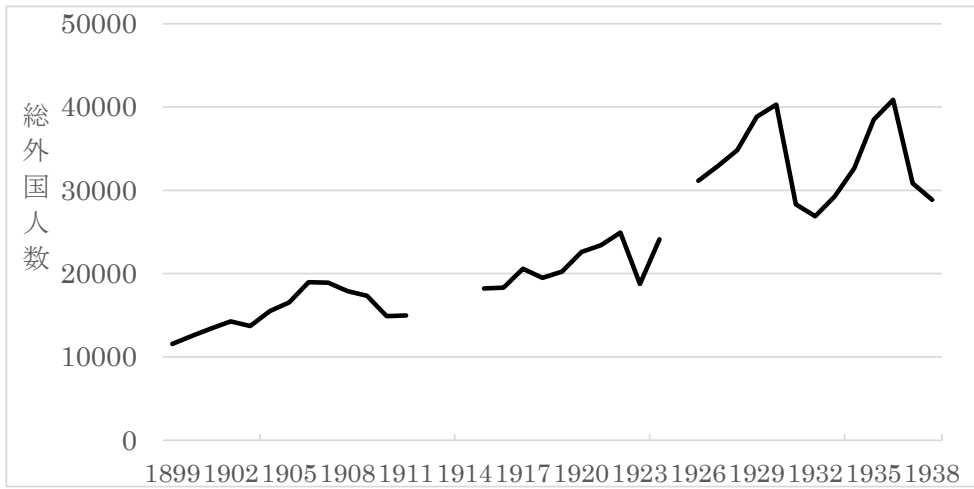
所得税額については、制度として整備されたのが1902年であり、それ以降のデータとな

っている。所得税額は県ごとの生産高を分析するために収集している。本来であれば県民所得が望ましいが、県民所得の年別データは存在しないようであり、そのため代替変数として所得税を収集した。今回収集したのは所得税の課税対象となる県別所得総額に対し、税率をかけたもの、つまり実際に徴収された金額となる。しかしながらこれには二つの問題があり、一つは各年で税率が異なっており、経年比較に適さないことである（詳しくは大蔵省主税局、1988 参照）。二つ目の問題として、当時の所得税は現在でいう法人税、公債・社債利子、そして所得税の三つを混合したものとなっている（それぞれ第一種、第二種、第三種所得税と呼ばれていた）。そのため総額では所得を十分に反映したものにはならない。第三種所得税の対象、すなわち課税対象となる個人所得の県別総額を収集するほうがより適していると言えるが、こちらにも問題がある。所得税の課税対象は一定所得以上を稼いでいるものに限定され（300 円以上など）、その基準となる所得も何度か変動している（詳しくは大蔵省主税局、1988 参照）。加えて、課税対象となる所得のデータが収集され始めたのは 1903 年からであり、こちらのデータを用いると 1 年分のデータを捨てることとなる。ただ課税対象となる納税者人数もわかっており、一人あたり課税対象所得を計算できるという意味では、こちらの変数のほうが望ましいかもしれない。

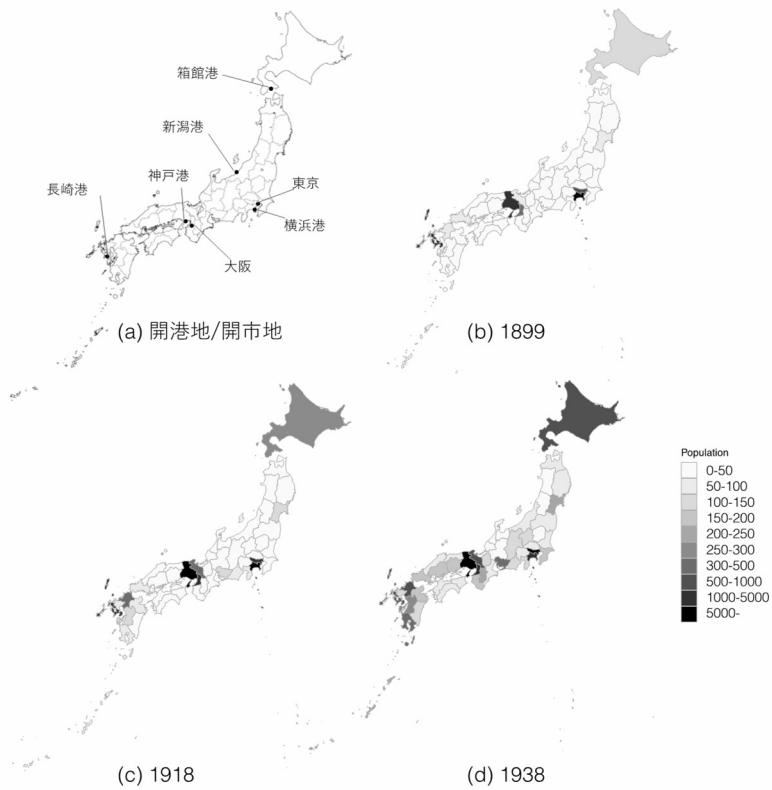
企業数についてだが、これ以上細かい分類はない。また本来であれば企業の生産高などを用いたいところだが、あいにく企業数もしくは資本金しかデータがなかった。後述するが、米は石高、甘藷・馬鈴薯（じゃがいもとさつまいも）は昭和 3 年までは貫で記載されており、それ以降はキロリットル（ヘクトリットル）で記載されていた。

これらの変数の中で、外国人数について少し掘り下げて論じるために、図を示す。図表 3 では年別総外国人数の推移を、図表 4 では日本地図を用いて外国人数の広まりを示している。1912 年、1914 年、1925 年はデータが記載されていないので空欄となっている。この図は大きく分けて 2 点の特徴があるといえる。1 点目は日本に居住する外国人数が右肩上がりの傾向にあることである。確かに年に応じて増減はしているものの、基本的には減少後すぐに回復し増加するというパターンをたどっているといえる。2 点目は大きな災害や恐慌の際には外国人数が減少しているということである。例えば図表 3 からは、1923 年の関東大震災や、1931 年以降の昭和恐慌、1937 年以降の日中戦争時に外国人数が大幅に減少している。こうした傾向は現在にも通じるものがあり、例えば 2009 年の金融危機や 2013 年の東日本大震災では外国人数は大幅に減少している。関東大震災後は復興特需ともいえるべきものが起きているのだろう、外国人数は V 字カーブを描いて増加している。恐慌後は時間を置いて回復しているが、これも現代に通じる傾向といえる。

次に、図表 4 に外国人数をプロットしたものを示す。図表 4a は開港地、開市地の場所を示したものである。開港地は箱館、新潟、横浜、神戸、長崎であり、開市は東京と大阪である。図表 4b から 4d は 1899 年、1918 年、1938 年それぞれの時点における外国人数を日本地図にプロットしたものである。開港地・開市地ではやはり外国人数が多いが、その中でもバリエーションがある。横浜や神戸は多いものの、北海道や新潟になるとその他の地域とそこまで変わらない。1899 年の時点では当然ながら開港地・開市地に外国人が集中しているが、年を追うごとに、特に 1938 年になると全国に外国人が散在している様子がわかる。では、こうした移住過程はどのように説明できるのだろうか。現在論文としてまとめているところであり、今後はそちらの論文にて結果を報告する予定である。



図表 3 : 1899 年から 1938 年にかけての総外国人数



図表 4 : 1899 年から 1938 年の 3 時点における外国人数プロットと開港・開市場の位置

III 実際に電子化しての雑感

本節では上述したデータを電子化した際の雑感について簡潔に述べる。まずもっとも関心の高いと思われる、どの程度時間を要するかという点である。小松・宮澤両氏の入力作業時間に関する情報も掲載することを考えたが、お二人に入力してもらったのはあくまで必要な変数の抜き書きであるため、1 ページ当たりにかかる時間の計算が少々ややこしいことになる。そこで改めて自分で 1 ページ分 (15 行×8 列プラス変数名) を入力してみたところ、9 分半で入力できた。巻ごとに揺れはあるものの、入力の際に使った年鑑は約 1200 ページであるため、1 冊を入力し終えるのに 190 時間かかる計算となる。また今回入力した情報は行数も列数も比較的少ないものであるため、これ以上時間を要すると見てよいだろう。スキャンデータと OCR を活用して取り込むという可能性もある。しかし紙自体が非常に薄く、裏写りが激しいため、十分に質の高い機材が必要となるだろう。

データとして入力する際にも、他にも問題点がある。1) 形式、2) 欠損値、3) 一貫性、の観点からまとめる。

1. 形式

『日本帝国統計年鑑』の形式は現代の統計表の形式と大幅に異なっているため、入力 (手入力、機械入力にかかわらず) 大きな困難が予想される。まず第 1 巻 (1882 年、明治 15 年) から第 23 巻 (1904 年、明治 37 年) までは漢数字でデータが表記されている。これらの巻はすべて縦書きであるため、右開きである。アラビア数字に慣れている分、入力作業においては多少困難となるかもしれない。地の文も漢字カタカナ混じりである。また都道府県の順番も異なっている。官庁データを利用したことのある人ならわかると思うが、都道府県の並びは北海道を 1 番、沖縄を 47 番とした並びで統一されている。しかしながら『日本帝国統計年鑑』では、第 1 巻から第 9 巻までは東京からはじまり京都、大阪、神奈川、兵庫、長崎といった都市圏の順になっており、第 1 巻では都道府県の数も 40 である。第 10 巻から第 32 巻までは東京、神奈川、埼玉、千葉、といった地域ごとの順になっている。第 33 巻から第 42 巻までは北海道からはじまる現在の都道府県の並びに近いものとなるが、宮城県などの位置が微妙に異なっている。第 43 巻からは現在の並びとなっている。あくまでこれらは慣れの問題であるものの、順応するまで多少時間を必要とするだろう。

またその他の形式の問題点として、非常に複雑な表の作り方をしている。図表 5 に明治 37 年の獣医の数を示したデータを例示したが、変数が何重にも入れ子になっていることがわかる。例えば「獣医」の下に「免状下付人員」と「仮免状下付人員」があり、さらに「免状下付人員」は 3 つのカテゴリに分か

部中州本		地方		
神奈川	東京	及試験	免状下付人員	獣医
一一	四一	校官卒立業学		
一	五八	校府卒県立学		
一三一	二四六	計		
一六	一二九			
四	七	付仮人員免状下		
二〇	一三六	合計		

図表 5：第 23 回日本帝国統計年鑑の獣医数に関する表

れている。加えて、こういった表では一セル一情報が原則であるが、なぜか「府県立学校」の項には二つの情報が入っている。この情報については特に説明がない。これ以外にも「0」は四捨五入の結果1未満になったものを示し、「|」は該当数なし（実質的に0）を示し、「…」は不詳、などといった独特の記法を用いている。これに加えて、ある年では不詳の項目それ自体を書かないことで示す場合もある（例えば外国人のいない都道府県は、表に記載されない）。電子化の際にはこうした困難を乗り越える必要があるだろう。

最後に、各年鑑は「明治〇〇年刊行」などと表紙に記載されているが、この年が必ずしも各種統計値が収集された年と同一であるとはいえない。例えば第29回日本帝国統計年鑑は明治43年に刊行されているが、「現在本籍人口」は明治41年の値、「本邦在留外国人」は明治42年の値、「国有鉄道職員数」は明治43年の値、といった具合である。これは一機関が統一的にデータを収集するのではなく、各省庁・機関が収集したデータを一冊にまとめるという『日本帝国統計年鑑』の性格からくるものと思われるが、それでもなおこの点に関しては自覚的に利用すべきだろう。

2. 欠損値

欠損値自体は避けられない問題だが、公的統計として欠損値が存在するのは比較的珍しいのではないと思われる。例えば1899年の外国人に関する表では、「データが期日までに届かなかった」という理由で京都の値が欠損となっている。理由は様々だが、ある年の一地域のデータが欠損となっているケースは『日本帝国統計年鑑』において比較的多く見られる。

また、ある変数の値が年によってまるまる収集・収録されていない場合もある。これは制度的な理由と、非制度的な理由の二つから生じている。制度的な理由として、例えば所得税という制度が導入されていなかった1902年以前には、当然所得税という項目は存在しない。また外国人が決められた地点以外に居住することが許されなかった時代（1899年以前）には、決められた地点以外における外国人数というデータは存在しない。他方、こうした制度的な理由とは関係なく、ある年の変数が収集されていない場合がある。例えば図表2で示したような3つの変数（外国人数、犯罪者検挙数、製糸工場数）はそれぞれある年度のデータがそっくり欠損になっている。

このような単年の欠損であればよいが、十数年にわたって収集していたにもかかわらず、それ以降は収集しなくなる変数もある。例えば県別の石炭消費量は、都道府県別の経済とエネルギー消費に関して重要な知見を与えてくれる変数と言える。しかしながらこの項目は第9巻から第25巻まで継続的に収集されていたにもかかわらず、第26巻以降は日本全体の値を収集するにとどまっており、その後第33巻以降は年次データすら収集されなくなっている。こうした十年単位の欠損は分析の幅を狭めるため、自分の関心のある変数がどの程度の期間にわたって収集されているかを事前に確認する必要があるだろう。

3. 一貫性

年ごとに一貫性のない収集をしている場合がある。一貫性のなさは、同一の内容を異なる項目名で指している、同一の項目名だが年ごとに内容が異なる、という二点にまとめられるだろう。

まず、同一の内容を指していると思われるが、項目名が異なっている場合がある。例えば「検挙された犯罪件数（犯罪者数）」という項目は年によって「全国犯罪者捕拿及自主」「就捕犯罪者違警罪犯罪諸規則違反者」「道府県検挙犯罪人及警察犯慮罰令諸犯則人員」「司法警察官ノ取扱ヒタル犯罪者及犯罪嫌疑者」「犯罪検挙件数」など、様々な呼称で呼ばれている。対象となる犯罪の種類はおおまかに重複しているので同一項目と捉えて差し支えないと思われるが、利用する際には注意が必要だろう。

次に、同一の項目名だが内容が異なる場合について記述する。先述のように、日本式の度量衡（石や貫など）から西洋式の度量衡（キログラムなど）に昭和4年頃から変換される。変換式自体は一貫しているものの、分析や記載の際には注意が必要だろう。第3章で収集した県別外国人数だが、これも内容が年によって異なっている。具体的には、1910年の韓国併合以降は韓国人が外国人数に入らなくなっている。これは制度としては一貫性をもっているが、経年比較の際には考慮すべき事項だろう。所得税も同様の問題を抱えている。各年で課税免除額の下限が異なっており（1887年以降300円以上、1913年400円以上、1918年以降500円以上、1920年以降800円以上、1926年以降1200円以上、1938年以降1300円以上）、比較には限界がある。

IV その他有用な戦前のデータ

本稿では『日本帝国統計年鑑』について紹介した。しかしながら戦前のデータはこれに限らない。他のデータと組み合わせることで、よりロバストな分析となるだろう。例えば物価は年によって大きく異なっているといえる。戦前の物価水準に関しては（1901年以降ではあるものの）日本銀行がデータを提供しており（参考文献参照）、これを用いることで賃金や流通額の物価を補正することができる。県別ではないものの、戦前の経済に関する包括的な研究を一橋の研究グループが行っており、その成果は『長期経済統計』としてまとまっている（大川・篠原・梅村、1965）。また当時の天気については年鑑降雨量の合計値や日照時間などのデータが気象庁により提供されている（なぜか奈良県だけ戦前のデータがないが）。第2章で示した研究関心に照らし合わせると、天気のよい地域に外国人が移住するという仮説の検定を行える。また天気の情報を操作変数に用いることもできるだろう。

V まとめ

本稿では『日本帝国統計年鑑』の概要、使用例、データ収集の際の問題点について述べた。『日本帝国統計年鑑』は未だに電子化が進んでいないが、重要性は述べたとおりであり、戦前の日本社会に関する新たな知見を多く与えてくれる。特に経済学や社会学において計量手法を通じた歴史研究が盛んになっている昨今において、同データは大きな意義を持つだろう。本稿で述べた使用例では一部のデータしか収集できなかったが、今後は統計年鑑全体を電子化することで、より柔軟な統計分析をのぞむことができるだろう。

参考文献

- 大川一司・篠原三代平・梅村又次監修, 1965, 『長期経済統計』東洋経済新報社.
- Grossman, R. S., & Imai, M, 2008, The evolution of a national banking market in pre-war Japan. *Explorations in Economic History*, 45(1): 17-29.
- 日本銀行, 2019, 「戦前基準の物価指数」(2020年1月7日アクセス) <https://www.stat.go.jp/library/faq/faq17/faq17a07.html>.
- 大蔵省主税, 1988, 『所得税百年史』, 大蔵省主税局.
- Tang, J. P, 2014, Railroad expansion and industrialization: evidence from Meiji Japan. *The Journal of Economic History*, 74(3): 863-886.

Summary

Digitizing “Statistical Yearbook of the Japanese Empire”

Akira Igarashi

This study introduces the characteristics and importance of “Statistical Yearbook of the Japanese Empire”, a series of statistics collected annually from 1882 to 1940. In this study I describe the experience of digitizing a part of the book, difficulties faced during digitalization, future usage of the book, and useful datasets from the pre-war period. For the purpose of my research, I digitalized information regarding the number of foreigners in each prefecture as well as 13 other variables, obtaining insights for foreigner situations during pre-war Japan. However, I experienced three types of problems during the digitalization including the format of the book, missing variables, and consistency of variables across years.

Keyword: Statistical Yearbook of the Japanese Empire; pre-war; number of foreigners

《研究ノート》

質的データアーカイブ構想の現状と課題

数値化されていない調査データの保存と活用に向けて

高橋かおり

【要旨】 本稿では、社会科学の質的データアーカイブの可能性について、数値化されていないデータの保存の観点から国内外の議論と事例の整理を行った。日本では2010年代に入り、インタビューデータを中心に質的データの保存や混合手法の活用などが議論されている。他方イギリスでは半世紀以上前から、フィンランドでも21世紀初頭からすでにデータアーカイブでの質的データの受け入れを始めている。そして各機関が仔細な基準やルールを公開しており、支援体制の充実は日本との比にならない。今後は調査現場とアーカイブ側で断絶している議論を架橋するプラットフォームや協力体制の構築が急務であると同時に、ボトムアップの運用方法の改善も求められよう。

キーワード：質的調査，データアーカイブ，調査遺産

I はじめに

本稿では、社会科学のデータアーカイブ構想と構築に関する議論の整理を通じて、数字にできないデータの扱いについての現状と課題を考察する。

RUDAをはじめとする、社会調査のデータアーカイブの機能はマイクロ（ミクロ・個票）データの収集、整理、保存、提供にある。加えて、実証研究の「再現性」を担保し、データを社会科学研究の共有財産とすることを推進する機構でもある（佐藤 2012:11）。東京大学SSJDAの始動が1998年であり、その間に2008年の社会調査協会の設立などを経て、この四半世紀で日本の社会調査をめぐる状況は大きく変化してきた。その中で、社会調査データアーカイブの構想・運用開始時点では想定していなかった、あるいは対応しきれなかった問題も生じている。

そこで、本稿では数値化されていない調査データという言葉を用い、インタビューや文書、写真や映像などの質的調査データのアーカイブの構想と運用について、国内外の取り組みを紹介する。本稿で扱う質的データや歴史的調査データは、データアーカイブの基本理念である実証研究の「再現性」の保証について、課題を残している。その課題や展望を整理したうえで、RUDAの今後の運営に向けての指針を示す。

II 国内における試みと議論——調査遺産（リサーチ・ヘリテージ）の保存

数値化されていないデータは、インタビュー記録（音声・スクリプト）やフィールドノー

ツ、メモや資料文書といった言語情報に限らず、写真や映像、図画や地図なども考えられる。映像や画像などのイメージを用いた社会学の質的調査の実践と教育は、言語情報の扱いに比べれば相対的に遅れている¹⁾。そのこと自体もまた問題ではあるが、本稿では先行議論の蓄積と筆者の専門性を踏まえ、言語化された質的データのアーカイブ化に焦点を当てる。

1. インタビューデータの扱い

まず国内における質的データアーカイブに関する議論を見ていきたい。質的データアーカイブについての実際としては2011年から小林多寿子らが行った調査の報告書が現状を把握するのに有用である(質的データ・アーカイヴ化研究会 2014)。本報告書では質的調査を行う研究者たちへの質問紙調査²⁾、日本社会学会と関西社会学会でのシンポジウムならびにダニエル・ベルトーによる講演会の記録、そして森岡清美の調査に関して本人の聞き取りを交えた整理が行われている。

小林(2014)は、この質問紙調査の結果を踏まえ、質的データアーカイブを設立するために必要な調査データの公共性の認識を訴えた。例えば、今日インタビュー調査において調査者は、インタビュー許可と同時に、インタビュー内容の利用許可——内容の確認・修正・削除、あるいは固有名詞の仮名・匿名の扱いなど——を得、データの収集と分析利用の2段階の許諾プロセスを経る。これに加え、アーカイブに加えることとなればアーカイブ所蔵のための許諾も必要になる。このことは、調査者に「自らの調査で得たデータをいかに考えるのか」が問われる。さらに公的資金を用いた調査であれば、「公共財としての調査データ」という認識はより強くならざるを得ない。

他方、データの貴重性や重要性、あるいは公共性だけではなく、調査対象者の意思の尊重や秘匿性も同時に問題となる。量的調査よりも個人の内面に迫った情報を多く聞き取るため、調査倫理の観点から調査者以外の第三者にデータが渡ることには慎重にならなければならない。さらに、質的調査においてはインタビュー以外にもフィールドワークや資料分析の組み合わせによって考察されることも多く、インタビューデータだけでは調査情報が完結しない場合がある。

この点について桜井厚(2003)は、社会調査においてラポール概念が重視され、調査者と協力者の間で構築的に語りが生産されるようになった経緯と背景を踏まえ、「警察の尋問みたい」と調査対象者に言われた経験を述べた上で「調査対象者を『他者』化することなく、また調査者の過剰な『自己』への思い入れにも陥らない社会調査とその方法論が求められている」と今後の展望を描いた。しかし、調査者という個人と被調査者という個人の相互作用から生まれた語りを、アーカイブという公的な空間に置くためには、個人の裁量や個人の技能で判断される調査対象者とのラポールを、社会科学の調査一般に対する信頼へと変えていく必要がある。

そしてこのように調査全般に対しての信頼を社会的に得るために、小林はアーカイブールの確立とアーカイブシステムの設立を求めている。具体的には、「調査対象者への許諾」や「公開手続きの明確化」のみならず、「非公開期間の設定」や「公開対象の限定」などといった点でも手続きの方針を確かにすることが求められる。

公開手続きの明確化や方針は RUDA にも存在する(朝岡・前田 2015:51-54)。しかし、質的データについて特筆すべきは、保存はするが公開しないという選択肢を積極的に取る

場合があることである。歴史的・学問的な価値はあったとしても即時公開するには本人への影響がある場合、数年、あるいは数十年と公開を留保したり、データへのアクセス要件をより厳しくしたりすることも求められよう。果たしてこのような運用が現実的に可能であるかは、制度設計を含めて考えなければならない。

2. 大規模質問紙調査とインタビュー調査の組み合わせ

日本家族社会学会全国家族調査委員会が1988年より定期的を実施してきた「全国家族調査(NFRJ)」は、2018年度調査で大規模質問紙調査に加えて半構造化面接によるインタビューと参与観察という2つの質的調査を併用した調査計画を構想し、2019年には実際の段階に移っている³⁾。調査委員会メンバーの1人である木戸功はインタビュー調査の実施について次のような期待を持っていた。少し長くなるが以下に引用する。

多くの場合、調査のプロセスと切り離すことができない質的データを量的データと全く同じように公開可能なものとみなすことはできないが、それでも質的データなりの公開の可能性を検討し、それを実現した際には、データの二次利用の可能性だけではなく、事後的な検証の可能性を担保することにもなる。個人研究として実施されるような質的調査においても、今後は分析の質を担保するためにデータの公開が求められるようになっていくとするならば、NFRJへの参画は個々の研究者にとっても有用な経験となるだろう。またそうした調査の取り組みは、質的調査それ自体の社会的理解を高めていくことにもつながるかもしれない。筆者は全国家族調査という観点、つまり現代家族の実態解明という観点のみならず、社会調査という観点からもこの取り組みに関心を持っている(木戸2016:219)。

ここで木戸が述べるNFRJの調査としての可能性は、データのアーカイブ化が果たしうる二次利用と事後的な検証(再現性)の役割を踏まえている。質的データアーカイブの難しさを前提としつつも、調査票と組み合わせで分析するために質的データの保管と活用を目指している。木戸は、小林(2014)や桜井(2003)の議論も踏まえて公開に向けた慎重さも持ちつつも、調査時点からアーカイブ化を見据えた調査設計をしていることを強調する(木戸2016:222)。さらに、調査グループでの勉強会的な研究会を重ねることで、個人の技や個々のラポールに回収されない、共有知としての調査手法の共有を行おうとしている。今後他にも同様の混合手法による調査がことも予想され、そのためにはデータアーカイブ側も受け入れ態勢や方針を整えることが求められる。

Ⅲ 質的データアーカイブ運用の実例——イギリスとフィンランドを例にして

国内では社会科学の質的データを扱ったデータアーカイブはまだ運用に至っていないが、海外ではいくつか実際に稼働している。本章ではそれらのデータアーカイブが公開している寄託ガイドラインを参照し、データアーカイブへの寄託に適した質的データについて、インタビューのトランスクリプションを例にみていきたい。

1. UK Data Service の方針

UK Data Service は, Economic and Social Data Service (ESDS), Secure Data Service, そして the Census Programme が統合して 2012 年 10 月にできた機関である。筆頭組織は 1967 年より稼働しているエセックス大学の UK Data Archive であり, そのほかに 6 つの大学と団体が運営に関わっている。UK Data Service は量・質両方のデータを扱っている。データの管理に際しては詳細な運用方針が web で公開されている⁴⁾。

インタビューのトランスクリプションの寄託についての指針を示したウェブページでは, 寄託に適したトランスクリプションの要素として, おもに形式面での条件をあげている。仮名や番号によってそれぞれの識別できること, 調査ごとにレイアウトが統一されていること, 話者表記が適切であることといった文書形式にとどまらず, 調査詳細が書かれた依頼文もそろっていることなども条件としてあげられている。そして倫理面では匿名化や仮名化も条件となる。

寄託の準備段階としては, 文書内の書式の統一やスクリプトの表記の説明, あるいは非英語話者のための要約の作成など, 単なるインタビューの書き起こしにとどまらず, 第三者でも理解できるようなメタデータやガイドラインも作成する。これは再利用する際に不可欠な情報である。あるいは, インタビューデータに限らず寄託データについてはファイル名やフォルダの階層化も秩序だったものにしなければならない。

メタデータをそろえることは量的データセットでも共通事項であるが, 質的データの場合, メタデータの指標が一律ではない場合が多い。アーカイブのガイドラインを考える場合も, それぞれのトランスクリプションの粒度(会話分析用なのか, 内容が分かればいいのか)によって異なる。質的調査データの場合, どのような分析目的で得られたのかということは, 最終的なトランスクリプション形式を大きく左右するのであり, 二次利用する人に対してその情報を提供することは必須であろう。

このような UK Data Service の方針の根底にあるのは, データ・ライフサイクルという考え方である。この考え方は近年の IASSIST でも主要なテーマの 1 つである(朝岡・高橋 2019)。すなわち, データを収集と分析にとどめるのではなく, 調査の設計, データの収集, データの処理と分析, 成果の出版とデータの共有, データの保存, データの再利用, そしてさらなる調査の設計というようなサイクルを調査データはたどるのが望ましいというものである⁵⁾。これは海外のデータアーカイブで共有される運営の基盤である。データは活用されなければならないのであり, そのサイクルに乗せるためには, 質的データも量的データと同様にデータ形式やメタデータの規格化が必須なのである。

2. Finnish Social Science Data Archive の試行錯誤と手厚い対応

次に, フィンランドのデータアーカイブの取り組みを見てみよう。1999 年に設立された Finnish Social Science Data Archive (FSD) はその設立初期から社会学者の Arja Kuula-Luumi が質的データアーカイブ運営に関する報告や論稿を発表しており, 準備段階から萌芽期を経て定着するに至るまでの思考錯誤や困難をたどることが可能である。以下では質的データアーカイブの設立から運用に至るまでの諸事例や諸問題をみていきたい(Kuula 2000; 2011, Bishop and Kuula-Luumi 2017)。

FSD が質的データの受け入れを始めたのは 2003 年である。その背景には社会科学における、量的調査方法に対抗した質的調査の再興がある。日本ではこの流れを受けて生活史研究が盛り上がりを見せるが、フィンランドではデータアーカイブにおける質的データの活用へと舵を切っている。つまり、個々人のデータ資源をどのように活用するのかということが焦点となった。そして、質的データアーカイブの整備はインフォームドコンセント（同意書への署名など）や調査倫理を研究者同士が共有することによってはじめて成立するのである。

現在 FSD では Alia Data Service からオンラインでデータをダウンロードすることができる。フィンランド語と英語に対応しており、フィンランド国外からの登録者は 10% にのぼる⁶⁾。また、質的データセットは 1476 件中 225 件である（2019 年 12 月 29 日現在）。

データ検索において特筆すべきは、ファイルの対応言語（英語・フィンランド語・スウェーデン語）、データの種類（質的・量的）以外に、データセットの利用方法について以下の 4 段階設定がされていることである。

- (A) 登録なしでどのユーザーも利用可能
(利用都度メールアドレスと目的を入力するだけで利用可能。ただし学部生は指導教員のメールアドレスを併記する必要がある)
- (B) 研究・教育目的での利用可能
- (C) (院生を含めて) 研究目的のみで利用可能
- (D) データの寄託者・製作者の許可を得た人のみ利用可能⁷⁾

このような指標は、UK Data Archive はもちろんのこと、SSJDA や RUDA においても設置されている。SSJDA は教育目的か否かでの検索が可能であり、例えば JGSS の利用は利用者リストの提出後に利用が可能になるなど、これらの基準は暗に設置されている。しかし、時点でわかりやすく明示されていることは利用者にとって利便性が高い。例えば (A) の都度ごとの情報入力で利用可能なデータセットが 1476 件中 85 存在する（2019 年 12 月 29 日現在）。さらに、英語版がないデータセットの翻訳サービスも存在する。

質的データの寄託については、細かくウェブサイトを紹介がある⁸⁾。これは、UK Data Service の指針と基本的には変わらないが、より詳細に質的データを寄託するための注意点が示されている。以下では UK Data Service では示されていない、FSD がより明確化・明文化した指針について見ていきたい。

まず FSD は、インタビューの粒度について、細かければ細かいほど良いという指針を出している。具体的には要約 (Gisted/summary transcription)、基本レベルの (ケバどりした) トランスクリプション (Basic level transcription)、正確な (素起こしの) トランスクリプション (Exact transcription)、そして会話分析用のトランスクリプション (Conversation analysis transcription) の 4 段階を示す。再利用のために寄託するのであれば要約のみは適さないとはいえ、細部まで書き起こすかは調査対象者の方針や利用可能な資源によるところが大きい。

さらに量的・質的にとどまらず個人データの匿名化については詳細なガイドラインがあり、寄託者と相談しながら二次利用に向けたデータセットの作成の手順がとられる⁹⁾。この

ようにデータアーカイブ側がデータの種類や個人情報に関する基準を設けることで、その基準が研究者の間に広まる役目を果たしている。

IV 質的データアーカイブを利用した研究——武田尚子のイギリス貧困調査を例に

イギリスやフィンランドなどにおいて質的データのアーカイブは実際に運用されており、利用者は全世界に開かれている。しかし日本語で読める狭義の質的データアーカイブの二次分析に関する研究は武田尚子の諸研究に限定されるだろう。前章ではアーカイブ側から見た質的データ利用を論じてきたが、以下では武田(2009:204-229)による実際の利用過程をもとにして、利用者側から見た質的データアーカイブの役割と日本での運用可能性を考えていきたい。ただし、武田の調査時期が2000年代中ごろであるため情報の更新は必要であり、イギリスのデータアーカイブの組織改編や、技術環境の変化を考慮して読み解かねばならない。

武田の調査記録からわかることは、社会科学のデータアーカイブの資料分析であっても、統計的処理を目的としたデータアーカイブが提供するデータセットの分析と、質的分析のためのデータセットの分析は全く異なることである。そのため、調査にかかる期間も労力も、専門職員とのコミュニケーションの量も頻度も格段に異なる。これは歴史学の資料調査のやり方に類似している。

第一に、オンラインでダウンロードできるデータもあるものの、現地に行かないと閲覧することのできない調査資料が圧倒的に多い。質的データアーカイブは、リサーチ・ヘリテージとして複数の調査にまつわる資料や情報を保管しており、電子化されていない、電子化できないものも多数ある。そしてこれらの「先人たちが調査にかけた苦闘のあとにふれることによって、現役の研究者が自らの調査方法や技術、モチベーションを高める貴重な手がかかる」のである(武田 2015:70)。つまり、調査分析の再現や検証にとどまらず、二次分析を通じて新たな知見を得るためには、調査結果のみならず調査過程から見直していくことが不可欠なのである。

しかしそこには問題もある。今日ではデジタル化が進んでいるためある程度の進展や改善があるとは考えられるが、在外研究や長期の調査出張、あるいは留学などを通じて、現地のアーカイブに物理的に行くことが武田の研究計画の前提にある。この点は、メタデータとデータセット、調査票や報告書などが電子ファイルで完結する数量的データセットの分析とは勝手が異なる。

第二に、かかる時間と期間の長さである。武田のスケジュールでは、調査の計画から分析の終了まで約3年かかったことがわかる。その間、初めの面談、4週間のデータ現物確認が2回、その上で1年間の在外研究を行っている(武田 2009:213)。調査目的に応じてこの期間やプロセスは変動するものの、所蔵調査リストから候補を選び、現物を確認し、現地職員と相談しながら分析の資料を決めることになる。中には詳細にリスト化されていない資料もあり、現地での現物確認は必須である。

つまり、質的データの二次分析は、フィールドワークと同様に調査設計と資料との対峙を繰り返し、問いと目的を常に更新しながら調査と分析を進めていくことが求められる。

武田の調査過程の記録からは、量的調査データを前提とした既存の社会科学データアーカイブでは想定していない問題や示唆が含まれている。その中で最も大きいのは、単に事務手続きを行うだけではなく、所蔵データの性質・種類・特徴、そして資料を熟知した専門職員（データライブラリアン）の存在であろう。日本のデータアーカイブは計量的・統計的手法を用いてきた研究者や専門家が数多く関わっているが、将来的に数値化されていない情報を扱う場合は、また別の専門家が必要である。国内での運用を考えれば、例えば図書館の司書資格保持者によって補える面もある。しかし保存の手続きをするだけではなく、利用者と相談しながら資料群のあたりをつけていかなければならず、そこにはまた別の高度な専門知識とコミュニケーション能力が求められるのである。

V 課題と展望——データアーカイブ側の受け入れ整備に向けて

本稿では、国内の研究者の動き、そしてイギリスとフィンランドのデータアーカイブの事例をもとにして、質的データの二次利用の可能性について論じてきた。最後に、データアーカイブ側から見た数値化されていないデータの活用について見ていきたい。

そもそも本稿のタイトルを「数値化されていない調査データ」としたのは、単にインタビューデータやテキストといった狭義の質的データの保存だけではなく、調査票原票やメモ、あるいは資料なども含めたリサーチ・ヘリテージの保存と活用を考えていくべきであるという立場にたつてのことである。例えば、調査票原票の復元と再分析（相澤ほか 2013）や、調査資料群の再検討（徳安 2019）ということも行われている中で、データの性質を量的と質的に厳格に区別をすることは不可能だろう。これまでの社会科学データアーカイブは SPSS や R といったソフトウェアで分析できる行列のデータセットと報告書・調査票を提供することに主眼が置かれていたが、それ以外の資料をいかに保存し、二次分析へと活用していくのかは避けては通れない課題である。

その際、保存と利用の基準を分けることは重要である。寄託されたすべての調査資料を公開することは不可能に近い。データアーカイブは、データフォーマットや寄託プロセスの標準化、メタデータの規格化や個人情報の扱いに関するルールを定めそれに当てはまるデータを受け入れてきた。リサーチ・ヘリテージという考えに則れば、調査データの種類に応じた柔軟な対応が求められる。例えば、最新の NFRJ は質的調査もそのデータセットに含むため、国内データアーカイブの既存の寄託方針では受け入れられないことが予想できる。その際、必ずしも即時的な公開を前提とした寄託ではなく、寄託受付とその後の保存方針、そして公開判断の問題を便宜上分けて対応することになろう。そして、それぞれの利用については、国内諸機関が領域横断的に検討できる場が必要となる。そこでは、社会科学だけではなくとりわけ歴史学を中心とした人文学の知見が生きてくることもあろう。

2019 年現在日本学術振興会が中心となり、「人文学・社会科学データインフラストラクチャー構築推進事業」が行われている（廣松 2019）。このようなマクロレベルでの規格統一やネットワーク化の動きと同時に、個々の研究者や研究資料と向き合い、その調査にあった寄託方法や公開基準を定めていくこともまた、データアーカイブの現場では求められている。今後の RUDA においては、マクロレベルの動きを注視しながらも、利用者や寄託者のニー

ズにこたえながら丁寧な運用をしていかねばならないだろう。

注

1) ビジュアルイメージのアーカイブについては石田佐恵子 (2014) に詳しい。また、欧米の実践の紹介については Knowles and Sweetman (2004=2012) が日本語で読める基礎的な文献である。

2) 対象者は日本オーラル・ヒストリー学会、生活史研究会、ライフストーリー研究会の会員である

3) ウェブサイトの回答者に向けたお知らせのページには「一部の方に、追加のインタビュー調査のお願いをお送りしております」との文言が書かれている。

<http://nfrj.org/for-respondents.htm> (2019年12月15日確認)

4) Transcription, UK Data Service (2019年12月16日確認)

<https://www.ukdataservice.ac.uk/manage-data/format/transcription.aspx>

5) Research data lifecycle, UK Data Service (2019年12月16日確認)

<https://www.ukdataservice.ac.uk/manage-data/lifecycle.aspx>

6) Services, Finnish Social Science Data Archive

<https://www.fsd.uta.fi/en/services/#aila-data-service> (2019年12月16日確認)

7) Data Catalog, Finnish Social Science Data Archive (2019年12月16日確認)

https://services.fsd.uta.fi/catalogue/index?lang=en&study_language=en

1

8) Processing Qualitative Data Files, Finnish Social Science Data Archive (2019年12月16日確認)

<https://www.fsd.uta.fi/aineistonhallinta/en/processing-qualitative-data-files.html>

9) Anonymisation and Personal Data, Finnish Social Science Data Archive (2019年12月16日確認)

<https://www.fsd.uta.fi/aineistonhallinta/en/anonymisation-and-identifiers.html>

参考文献

相澤真一, 小山裕, 鄭佳月, 2013, 「社会調査データの復元と計量歴史社会学の可能性——労働調査資料 (1945-1961) の復元を事例として」『ソシオロギス』 37:65-89.

朝岡誠, 前田豊, 2015, 「データアーカイブ事業の展望に関する一考察」『社会と調査』 1:49-62.

朝岡誠, 高橋かおり, 2019, 「海外データアーカイブの動向 2—IASSIST 年次大会の報告から」『社会と統計』 5:33-41.

Bishop, Libby and Arja Kuula-Luum, 2017, “Revisiting Qualitative Data Reuse: A Decade On,” *SAGE Open*, 7(1):1-15.

石田佐恵子, 2014 「映像アーカイブズと質的研究の展開」『フォーラム現代社会学』 13:133-143.

廣松毅, 2019, 「人文学・社会科学分野におけるデータインフラストラクチャー構築の推進」

『ESTRERA』308:2-7.

木戸功, 2016, 「NFRJ と質的研究——質的データの収集と分析および公開に向けて」『家族社会学研究』28(2):218-223.

Knowles, Caroline, and Paul Sweetman, 2004, *Picturing the social landscape: visual methods and the sociological imagination*, Routledge. (2012, 後藤範章監訳『ビジュアル調査法と社会学的想像力——社会風景をありありと描写する』ミネルヴァ書房.)

小林多寿子, 2014, 「質的調査データの公共性とアーカイヴ化の問題」『フォーラム現代社会学』13:114-124

Kuula, Arja, 2000, “Making Qualitative Research Material Reusable: Case in Finland”, *IASSIST Quarterly*, 24:14-17.

———, 2011, “Methodological and Ethical Dilemmas of Archiving Qualitative Data”, *IASSIST Quarterly*, 34:12-17.

桜井厚, 2003, 「社会調査の困難——問題の所在をめぐって」『社会学評論』53 (3) 452-470.

佐藤博樹, 2012, 「実証研究におけるデータアーカイブの役割と課題——SSJ データアーカイブの活動実績を踏まえて」『フォーラム現代社会学』11:103-112.

質的データ・アーカイヴ化研究会, 2014, 『質的データ・アーカイヴ化とリサーチ・ヘリテージ』2011-2013 年度科学研究費成果報告書「質的データとしてのライフストーリーのアーカイヴ化と〈調査遺産〉継承の経験的研究」(課題番号 23530619).

武田尚子, 2009, 『質的調査データの二次分析——イギリスの格差拡大プロセスの分析視角』ハーベスト社.

———, 2015, 「質的調査データの二次分析——大正期『月島調査』と労働運動」『日本労働研究雑誌』665:70-80.

徳安慧一, 2019, 「森岡清美の調査資料からみる社会調査法の彫琢と社会調査教育——『業績にならなかった』調査資料群から」『一橋社会科学』11 (別冊):67-80.

Summary

The Vision and Mission of Qualitative Data Archives : The Preservation and Reusing of Non-numeral Data

Kaori Takahashi

There is no quantitative social science data archive in Japan. This paper considers the possibilities and difficulties of managing quantitative social data archives. In 2010, some researchers discussed the preservation and reusing of quantitative data and called it “research heritage.” In the UK, there are data archives that manage not only quantitative data but also qualitative data from over half a century. Also, Finnish Social Science Data Archive began to accept qualitative data too. These data archives have detailed rules and guidelines for users and contributors. It will be difficult to construct a quantitative social science data archive in Japan, but it is very much possible. For the first step, we need to recognize the value of research data as research heritage. Furthermore, it is essential to connect researchers and organizations.

Keywords: quantitative data, data archive, research heritage

《資料》

海外データアーカイブの動向 3

—IASSIST 年次大会の報告から—

五十嵐 彰
高橋 かおり

【要旨】 社会調査データは今後の社会の発展に寄与する公共財であり、広くデータが利活用される環境を整備する必要がある。本稿では 2019 年 5 月に開催された IASSIST 年次大会で報告された内容をもとに海外のデータアーカイブで行われている先進的な取り組みを紹介する。これらを踏まえ、今後の CSI 業務ならびに RUDA 運営に対しての方針と示唆を提案する。

キーワード：データアーカイブ，データ利用，データ共有

I はじめに

近年日本でもオープンサイエンスの動きが加速し、自然科学のみならず人文科学・社会科学におけるデータ利活用が政策的・学術的論点となりつつある。とりわけ国立情報学研究所（NII）が中心に進めているオープンデータサイエンス構想においては、各拠点間の連携を果たしつつ、共通してデータを共有・利用しあう基盤の構想が進められている（廣松 2019）。しかし日本におけるデータ利活用の議論とその動きは諸外国と比べると圧倒的に遅れているのが現実である。

CSI 社会調査部会では、社会科学に関する情報技術・データサービスに従事する専門家の国際組織である IASSIST が開催する年次大会に部会業務として毎年参加している。本稿では、前田・朝岡（2017）、朝岡・高橋（2019）に引き続き、IASSIST にて開催されたワークショップや研究報告をもとに海外のデータアーカイブで行われている先進的な取り組みを紹介し、RUDA 運営を含めた今後の CSI における業務の指針を提案する。

今回の年次大会のテーマは“Data down under”であり、オーストラリアの別称である Land down under をもじったものである。データはあらゆる意思決定の根幹にあるものであり、そうしたデータをいかに整備し利用を促すかといった側面に焦点を当てた発表が多く見られた。参加者は世界各国の大学教育部局スタッフ、ライブラリアン、メディアセンター職員などである。Business Meeting で示されたデータからは、大学のライブラリアンが占める割合が多いことが明らかになった。

本年度の研究大会では、全体に質問紙調査を保存する旧来的なデータアーカイブの役割を越えて、多様化するデータや調査の形式への対応や、データ・ライフ・サイクルの推進、寄託されたデータの活用やリテラシー教育、さらにはデータを利用した大学改革など、データ保存を前提として、いかにデータを活用するのかが議論の中心であった。さらに南半球での開催ということもあり、ヨーロッパ・北アメリカ中心のデータアーカイブの動

向ではない、オルタナティブな実践や試みが紹介されていた¹⁾。

II. 保存データの多様化への対応

1. データの質の維持

2018年にEU一般データ保護規則（GDPR）の適用が開始されたことにともない、各地のデータアーカイブもその対応が課題となっていた。とりわけ今回の開催国のオーストラリアでは、EUの規則を参考にしながらも、その問題点を改善しつつ、独自の保護規定を作る動きがある（Russell, 2019）。すべてのデータを公開に向けて保存するのではなく、オープンデータ化は推進しつつも、歴史的背景や文脈に即した利活用の方法と段階、時期の判断が求められる。

さらに、寄託されるデータの質の担保も複数のデータアーカイブにおいて運営上の課題となっていた。例えばUK Data Serviceではある程度自動的に寄託者がデータの質を確認できる「データ・ヘルス・チェック」と言われる制度が導入されつつあった（Corti, 2019）。寄託から公開までの過程を短縮する試みは、ICPSRでも行われている。これらの過程を一部自動化することは、アーカイブ側の時間や労力の省略に寄与するものの、すべての自動化は不可能であり最終的には専門家の判断が必要になるという見解は共通していた。

これらの取り組みから、データの種類が複雑になるにつれて、専門職者の果たす役割も増す一方で、寄託数の増加に伴い、利用者や寄託者側にも最低限のリテラシーや判断基準を持つことが求められていることがわかる。日本ではそもそも「データの寄託」という意識が共有されていない状況であるが、オープンデータ化を目指す前にデータを収集し保存すること、そしてそれを利活用するのはなぜなのか、という研究に関わる根本的な意識改革を進めることが不可欠であろう。

2. 多様な形式のデータ保存に向けて

データ利用者からの報告においては、既存のデータセットに収まらない形のデータ分析の例が示されている。ここには、旧来から存在する量的／質的という単純なデータの種類の問題だけではなく、技術の進歩によってこれまで取得できなかった、あるいは想像もしていなかった種類のデータの入手が可能になったことが要因にある。

例えば、画像データはその1つであろう。文化財保護の文脈において、写真や図面の保存に向けた取り組みの紹介があった（Kirby, 2019）。建物そのものだけではなく、建物に関するデータや保存は、万が一災害や事故などによって現物が破損や消失したとしても、修復や復元を可能にする助けとなる。ただし実際にデータを記録・保存するのであれば、そもそも何のために保存するのか、単なる搾取ではなく、コロニアリズムへの反省を踏まえた取り組みになっているのか、という問いかけもなされる。情報保存するという目的を定めるだけではなく、保存データに関しての著作権・肖像権などの法的・倫理的な利権関係も明らかにしておかねばならない。

あるいは、オンラインゲームのプレイデータといったビッグデータの収集も可能になっている。犯罪組織分析の代替モデルとして、犯罪をテーマにしたオンラインゲームデータの

分析をした事例においては、分析の可能性と同時に、IP 情報や利用者 ID をめぐる利用者のプライバシーの保護などが今後の展開に向けた課題としてあげられた (Fleet, 2019)。

さらに、社会科学データアーカイブの老舗である ICPSR においても、ソーシャルメディアの情報をどのように扱うのかということが議論されていた (Hemphill, 2019)。ソーシャルメディアの情報は可能性に満ちているとはいえ様々リスクがあり、質の良いデータを保存しつつ提供していくプラットフォームの構築には各所との調整が必要であるという報告がなされた。

新たなデータの保存・活用の事例からは、以下の点が課題として提出されている。第一に、データの保存の目的である。二次利用を目的とした保存であったとしても、ソーシャルメディアやゲームのように同時代の人に向けて活用促進を図るのか、あるいは文化財情報のように将来の人のためにに向けて保険として残すのかによっても保存のための制度設計は変わってくる。第二に、保存の際に生じる倫理的・法的手続きである。個人情報保護だけでなく、データの所有権の明確化も必要な手続きである。そして第三に、どのような形式で提供するかである。画像やネットワーク情報など、従来のデータセットよりも複雑なデータ形式になるのであれば、その利活用の際の説明やチュートリアルなどの充実も図らなければならない。

このように、取得できるデータが多様化する中で、これまでのデータアーカイブとしては形式が異なる資料やデータについても、その保存方法を検討していく柔軟な姿勢が求められるともいえよう。

III データ利活用の実際——研究以外での応用可能性

2019 年の年次大会においては、大学でのデータの利活用に関する教育的セミナー報告が複数見られた。例えばイーストミシガン大学の「Datathon」は、プロジェクトベースの市民向け学習プログラムであり、行政データの再分析や解釈を行いながら地域の水問題への理解を深める取り組みであった (Brodsky & Kelly, 2019)。参加者は高校生や高校教員などが多く、受講生がデータへの興味を持つ一助となっていた。プログラム運営においては実際にデータに触れて分析する時間の少なさや運営時における財政問題など指摘されていたが、これは大学の地域連携の新たな形の 1 つといえよう。

別の例では、学生が自学自習の過程で適切なデータにたどり着くための指針や援助のあり方が紹介された (Sheley & Arave, 2019)。インディアナ大学では CRAAP という指標を制作し、適切に流通しているか (Currency)、妥当性をもっているか (Relevance)、適切な機関や団体が発行しているか (Authority)、正しいか (Accuracy)、そして知りたいことの目的にあっているのか (Purpose) の 5 点を確認できるワークシートを学生に提供するようになった。

同様の例として、カールトン・カレッジでは「Data Squad」というデータ利活用に特化した学生スタッフを置き、おもに学部生からの質問に答えられるような制度を整えていた (Lackie, 2019)。

この 3 つの実践からは、データを利用できるようにするだけではなく、どのように利用す

ることが可能か、そしてその利用方法の適切さをいかに確認するのかということが、データ利活用教育において欠かせないということが明らかになる。CSI の各種セミナーやコンサルティングもこのような役割を果たしているといえるが、今後の展開や活動の広報をする際の参考になる事例であった。

さらに、アメリカのロチェスター大学では、データチーフという役職に抜擢された学外の専門家が、学内のデータ活用の実態とその散逸の状況を把握し、必要なデータを適切に提供する取り組みを行ったことで、大学改革につながったという報告があった (Cannon 2019)。この事例を立教大学に応用できるのであれば、データや資料を保有する複数の部局が互いの状況を知り、活用しあうことが、学術面のみならず大学のマネジメントの変化にもつながることが考えられる。

各機関のデータ利活用の取り組みからは、データ利用におけるコミュニケーションの受容さと、単に「データを分析する」という作業面だけではなく目的や活用例、そしてその影響も考えたマネジメントプランを考える思考を利用者側も分析者側も持つことの重要性が明らかになった。

IV 大学間ネットワークの構築

複数の報告で、大学間の協同ネットワークやその構築を志向していることが目立った。大学間ネットワーク構築の目的として、データアーカイブそのものの共有というよりは、アーカイブに関する様々な業務を共有して負担を減少させることを目的としているようであった。

例えば Data curation network (DCN) というデータキュレーション (データチェック・クリーニングなどの一連の過程) を行うネットワークを 9 つの機関で構築しているという報告があった (Blake & Herndon, 2019)。発表時の参加機関としてジョン・ホプキンス大学、デューク大学、ミネソタ大学、セントルイス・ワシントン大学、コーネル大学、ペンシルバニア州立大学の 8 つの大学、そしてデータリポジトリの DRYAD が挙げられていた。ネットワーク参加者の内訳として、20 人以上のデータキュレーター、7 人の代表、そしてアドミニストレータとプログラムディレクターとが一人ずつということであった。活動内容はキュレーションのステップの標準化、キュレーションツールやトレーニングデータの共有、ワークショップなどを通じたキュレーターの育成、ベストプラクティスの共有などであり、事例などを何本かの論文にまとめることもしている。

2 つ目の事例として、ICPSR のリサーチャーパスポートが挙げられる (Hemphill, 2019)。データアーカイブには、個人を特定できる情報などを含むデータに対して制限を加えており、こうしたデータが 1,482 件蓄積されている。制限データにアクセスするには、研究者個人がオンラインアプリケーションを行い、ICPSR の職員がそれを閲覧 (必要であれば修正の要求) し、アクセスを承認するという流れがある。このアプリケーションと承認のプロセスを簡略化するために、リサーチャーパスポートを提案している。リサーチャーパスポートはデータ利用者それぞれに与えられるアカウントで、データ利用者の情報 (博士号の有無、データ利用に関する受講の有無、利用経験など) を搭載している。リサーチャーパスポートはデータ利用者がどの程度信頼できるかを測るバロメーターとして機能し、データ利用の

許可の判断をリサーチャーパスポートが肩代わりすることができる。現在は ICPSR のみの適応となっているが、今後はデータアーカイブ間でパスポートを共有することも構想しているとのことだった。

以上のように、アーカイブ間の共同では、データの共有ではなくそれに関わる業務についての共有を行っていることが目立っていた。またこうした構想に共通しているのは、研究者ではなく、データキュレーターなどの技術者による先導だろう。技術者の専門的な知識と経験をもとにした主導により、こうしたネットワーキングや新たな構想が可能になるのではと考えられる。

V 再現性確保へのデータアーカイブの役割

社会科学の分野において、昨今再現性の問題が盛んに取り上げられている。トップジャーナルに掲載された心理学の研究のうち 39%しか再現に成功しておらず (Open Science Collaboration, 2015)、大きな議論を巻き起こした。これらは異なるデータを用いて同じ結果を示す反復可能性 (repeatability) を指しており、理論の形成と学問の発展にとって重要なものであるといえる (Freese & Peterson, 2017)。他方でより根本的な、研究者が用いたデータと手法で同一の結果を示す検証可能性 (verifiability) に関してはより整備が進んでおり、政治学のジャーナルでは論文採択時にデータを求めるものが多くなってきている。仮に検証可能性が確保されないと、最悪のケースでは研究者による捏造などを見過ごすことになる。こうした風潮に対応するために、データアーカイブが再現可能性 (検証可能性) を促すためのトレーニングなどのサービスを行っているという事例がいくつか散見された。

コーネル大学のデータセンター (CISER) では、再現性の確保のための新たな事業を立ち上げている (Heslop, 2019)。CISER では、投稿前の段階で研究者が実際に用いたコードとデータを用いて結果のダブルチェックを行い、さらにデータとコードをパッケージとしてまとめ、投稿後のデータ投稿要求に即座に答えられるように体制を整えている。利用者は 2018 年の段階では 5 件、2019 年では 20 件と順調に対象を拡大している。現在では研究者に限定されているが、今後は修士や博士の学生向けにサービスを拡大する計画のようである。

次に、イエールとコーネルから、Data CURE Training Program の紹介があった (Christian, Thompson, Peer, & Arguillas, 2019)。このプログラムは研究者に対する再現可能性に関するトレーニングを行うと同時に、情報専門職の職員に対してはデータマネジメントの方法に関する授業やワークショップを、リサーチスタッフに対してデータキュレーションに関するトレーニングを行っている。研究者に対するトレーニングで再現性に関する基本的な考え方を授ける一方で、再現性確保のために研究者以外のスタッフに対して、データの整備やコードの解読ができるように教育している。再現性の確保が大学の評判に関わるという考えからか、全学的な取り組みとして再現性確保のためのシステムが構築されつつあると断言していいだろう。ただこれはあくまでコーネルを中心とした取り組みであって、その他の大学についてどこまで当てはまるかは不明である。

IASSIST に関する事例ではないが、共著者の一人である五十嵐が過去に所属していたオランダのユトレヒト大学では、再現性に関する取り組みが 2010 年代初頭から行われていた。

2012年にオランダ人の社会心理学者であるディーデルク・スターペル (Diederik Stapel) による大規模な不正が発覚して以降、オランダでは再現性に慎重になっているといえる。ユトレヒトでは修士論文を含む全ての論文はデータとコード、そしてその目録をフォルダにまとめて大学に提出し、毎年ランダムにそれらのファイルを選んで再現するという取り組みが行われていた。こうした再現を行うのは研究者ではなくマネジメントスタッフであるため、今回のコーネルにおける取り組みのように、彼らに対するコードやデータに関するトレーニングは必須であったのだろうと思われる。

VI データ寄託の行動分析

最後に、アーカイブへのデータの寄託行動に関する分析を紹介する。データアーカイブにとってデータの寄託数を増やすことが1つの重要な懸念事項と言えるが、なぜ研究者はデータを寄託するのか（しないのか）という問いに対し、GESISの研究者が計量分析を行っていた (Akdeniz & Perry, 2019)。Ajzen (1991)の計画的行動理論をもとにし、寄託行動への態度、主観的な規範の認知、そして知覚された行動の統制可能性の三つの観点から寄託意図への効果を分析している。発表者達は2012年から2014年に社会科学のジャーナル (10の社会学、そしていくつかの政治学のジャーナル) から出版された論文の著者にオンラインサーベイを送付し、459人からの回答を得た。寄託行動への態度は3つの変数、規範認知は2つの変数、そして統制可能性は1つの変数によって測定されており、それぞれが寄託意図に与える効果を分析している。

寄託行動への態度は1) 他の研究者が新たな発見を行う可能性への態度、2) 自分の研究の評判が上がることへの態度、そして3) データの誤解釈の3つで測定されている。自分の研究への評判は寄託意図へ有意な正の関連、それ以外の2つは有意な負の関連をもっていた。次に主観的な規範として、データの寄託への期待とそれに応えること、そしてデータ寄託をするのがより一般的であるという考え方のそれぞれが寄託意図と正の関連があった。最後に、統制可能性を反映する、適切なデータリポジトリが利用可能だという尺度が寄託意図と正の関連をもっていることが明らかとなった。これらの結果に加えて発表者達は追加分析を行い、自分の研究やデータが認知される必要性という新たな変数が完全に規範を媒介していることを示し、データ認知の必要性に訴えること、そして必要性が低い場合は、インセンティブを与えることが重要であると議論していた。

研究者がデータを共有することはある意味で規範的な行動とも言えるが、他の研究者による新たな発見や誤解釈といったリスクも同様に認知し、寄託意図を妨げている。他方で自分の研究に対する評判や認知が向上することは寄託への誘引として働いているといえる。そのため、また外的なインセンティブも重要であり、この結果に従うと、例えばSSJDAの寄託者表彰などはデータの寄託を促進するといえる。

VII. まとめと今後への指針

本稿では IASSIST において紹介されていた取り組みをいくつか紹介した。リサーチャー

パスポートのように早急に取り入れることが難しいものも多いかもしれないが、例えばキュレーションに関するネットワークなどは大学間・アーカイブ間で取り組みをしても良いかもしれない。特に現段階ではアーカイブに関する共有の構想は盛り上がっているものの、その他の業務に関する共有は比較的下火だと思われる。また再現性に関する取り組みも、研究者に対するトレーニングに加えて、アーカイブが再現性確保のために協力することも将来的には構想することも可能だろう。

画像やゲームのプレイデータなど、従来データアーカイブが想定していた量的データとは異なる形式のデータも保存が可能であるということは、質的研究に大きく依拠している日本の社会学のアーカイブを考える上で大きな収穫であったと思われる。今後は量的データに拘らない、新たなデータアーカイブの形も模索することが求められるだろう。例えば本号の高橋による報告は、日本における質的データアーカイブの可能性についてまとめられている。こうした海外の非量的データアーカイブについての知見を参考にした、非量的データアーカイブを構築することにより、将来的に社会学のデータの保全に貢献することができだろう。

注

1) IASSIST 年次大会での発表資料は会議記録よりオンラインで入手が可能である。

<https://iassistdata.org/conferences> (2019 年 12 月 26 日取得)

なお、本大会における報告は(名字, 2019)で表記している。

参考文献

- Ajzen, Icek, 1991, The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
- 朝岡誠・高橋かおり, 2019, 「海外データアーカイブの動向 2—IASSIST 年次大会の報告から」『社会と統計』5:33-41.
- Freese, Jeremy, and Peterson, David, 2017, Replication in social science. *Annual Review of Sociology*, 43, 147-165.
- 廣松毅, 2019, 「日本学術振興会『人文学・社会科学データインフラストラクチャー構築推進事業』について」『ESTRELA』308:2-7.
- Open Science Collaboration, 2015, Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.
- 前田豊・朝岡誠, 2017, 「海外データアーカイブの動向—IASSIST 年次大会の報告から」『社会と統計』3:27-35.

Summary

The Trend in Foreign Data Archives 3 : From the Presentations at the IASSIST Annual Conference

Akira Igarashi
Kaori Takahashi

Social research data is a public good that contributes to the development of future societies, and thus we need to facilitate an environment in which people can use data appropriately. This paper reports on cutting-edge research and projects presented at the International Association for Social Science Information Service and Technology (IASSIST) held in May 2019. Based on these presentations, we suggest future directions for the Center for Statistics and Information (CSI) and Rikkyo University Data Archive (RUDA).

Key words: data archives, data usage, data sharing

2019 年度社会情報教育研究センター研究活動等報告

■ 政府統計部会 部会別研究活動等

1. 研究論文・著書

著書名：『日本の公的統計・統計調査』

発行機関：立教大学社会情報教育研究センター

発行年月：2019 年 10 月

著者氏名：櫻本 健・濱本 真一・西林 勝吾

共著区分：共著

櫻本 健 准教授

1. 学会発表

日時：2019 年 4 月 6 日

学会名等：国際価値論研究会

場所：立教大学池袋キャンパス 12 号館

発表者：櫻本 健

発表テーマ：「デジタルエコノミーの伸張による GDP に対する影響～地理的 GDP とアメリカでの研究を受けて」

日時：2019 年 9 月 6 日

学会名等：経済統計学会第 63 回全国研究大会

場所：東北学院大学土桶キャンパスホワイ記念館 3 階

発表者：櫻本 健

発表テーマ：「2030 年代に日本版インボイス制度によって SNA を改良できる技術的余地」

日時：2020 年 1 月 11 日

学会名等：経済統計学会東北・関東支部例会

場所：立教大学池袋キャンパス 12 号館

発表者：櫻本 健

発表テーマ：「多国籍企業の節税フレームを統計で捕捉する選択肢」

2. 研究論文・著書

著書名：『供給使用表と投入産出表に関するハンドブック（拡張と応用を含む）』（国際連合経済社会局統計部著の翻訳公表物。原典は次の通り。United Nations, Department of Economic and Social Affairs, Statistics Division, Editor Sanjiv Mahajan (2017) “The Handbook on Supply, Use and Input-Output Tables with Extensions and Applications”, https://unstats.un.org/unsd/nationalaccount/docs/SUT_IOT_HB_wc.pdf)

発行機関：発行せず、内閣府経済社会総合研究所「供給使用表と投入産出表に関するハンドブック（拡張と応用を含む）（仮訳）」ページに公表

http://www.esri.go.jp/jp/esri/gdpstat_kaizen/handbook/kariyaku.html

発行年月：2019年7月

著者：国際連合経済社会局統計部、編集者：サンジブ・マハジャン

日本語版監訳者：猪俣 哲史・櫻本 健・田原 慎二・萩野 覚・牧野 好洋

掲載種別：翻訳

共著区分：共著

著書名：「経済系のための情報活用1 Office2019 対応」

発行機関：実教出版

発行年月：2019年10月

著者氏名：櫻本 健（編著）・倉田 知秋・小澤 康裕・安藤 道人・菊地 進

掲載種別：著書

共著区分：共著

3. 依頼研修

日時：2019年6月12日、6月13日

研修名：「GDPを学ぶ」

主催：内閣府経済社会総合研究所経済研修所

場所：中央合同庁舎8号館416会議室

講師：櫻本 健

対象：内閣府職員、中央省庁職員

日時：2019年6月26日、6月27日

研修名：「国民経済計算(SNA)ステップアップ」

主催：内閣府経済社会総合研究所経済研修所

場所：中央合同庁舎8号館416会議室

講師：櫻本 健

対象：内閣府職員、中央省庁職員

日時：2019年6月24日、6月25日

研修名：総務省統計研究研修所2019年度統計専門課程「国民経済計算・県民経済計算」

主催・場所：総務省統計研究研修所

講師：櫻本 健

対象：中央省庁職員、自治体職員

日時：2019年8月27日、8月29日

研修名：総務省統計研究研修所2019年度統計専門課程「国民経済計算・県民経済計算」

主催・場所：総務省統計研究研修所

講師：櫻本 健

対象：中央省庁職員,自治体職員

日時：2019 年 10 月 11 日(金), 15 日(火),18 日(金),11 月 8 日(金)

研修名：本科（総合課程）「国民経済計算（サテライト勘定を含む）」

主催・場所：総務省統計研究研修所

講師：櫻本 健

対象：中央省庁職員,自治体職員

4. 依頼講演

日時：2020 年 1 月 31 日 午後 13 時 00 分～16 時 30 分

講演名：令和元年度県民経済計算担当者研究会議

主催：内閣府経済社会総合研究所国民経済計算部地域・特定勘定課

場所：合同庁舎 8 号館講堂

講師：櫻本 健

講演テーマ：「統計法改正と県民経済計算作成に及ぼす影響」

対象：内閣府経済社会総合研究所職員、自治体職員、テーマに関連した研究者

安藤道人 准教授

1. 研究論文・著書

掲載誌名：PLOS ONE, Vol 14, Issue 9

タイトル：「Mass Evacuation and Increases in Long-term Care Benefits: Lessons from the Fukushima Nuclear Disaster」

発行機関：PLOS

発行年月：2019 年 9 月

著者氏名（共著者含）：Tomohiro Morita, Michihito Ando, Yui Ohtsu

掲載種別：論文

共著区分：共著

掲載誌名：『立教経済学研究』, Vol.73, No.1, pp.59-68,

タイトル：難病政策の社会保障化における政策形成と利用者負担——2013 年度法案審議に関する考察——

発行機関：立教大学

発行年月：2019 年 9 月 7 月

著者氏名（共著者含）：渡部 沙織・安藤 道人

掲載種別：論文（紀要）

共著区分：共著

掲載誌名：『弱者を生まない社会へ ベーシック・サービスの実現をめざして』第 3 章, pp.53-75

タイトル：社会保障と教育における自己負担改革 「必要原則に基づく総合合算制度」案の検討

発行機関：公益財団法人 連合総合生活開発研究所

発行年月：2019年3月

著者氏名（共著者含）：安藤 道人

掲載種別：論文（報告書の一章）

共著区分：単著

掲載誌名：『月刊DIO 連合総研レポート』 No.348

タイトル：中間層と福祉国家：税・社会保険料・社会保障の再分配効果の諸研究の検討

発行機関：公益財団法人 連合総合生活開発研究所

発行年月：2019年6月

著者氏名（共著者含）：安藤 道人・古市 将人

掲載種別：論文

共著区分：共著

濱本 真一 助教

1. 学会発表

日時：2019年3月7日

学会名等：数理社会学会第67回大会

場所：立命館大学

発表者：濱本 真一

発表テーマ：「合理的選択による進学格差の変動モデル——相対リスク回避と教育拡大の関係の定式化」

日時：2019年8月31日

学会名等：数理社会学会第68回大会

場所：熊本県立大学

発表者：池田 岳大・濱本 真一

発表テーマ：「職業学科の地位達成過程に関する不安定性」

2. 研究論文・著書

掲載誌名：『教育社会学研究』104

タイトル：「隣接分野から見た高等教育研究への問題提起」

発行機関：日本教育社会学会

発行年月：2019年6月

著者氏名：相澤 真一・濱本 真一

掲載種別：論文（学術誌）

共著区分：共著

掲載誌名：『社会学年報』48

タイトル：「無業リスクに関する高校間格差とその趨勢——職業科トラックのセーフティネット機能の検討」

発行機関：東北社会学会

発行年月：2019 年 7 月

著者氏名：池田 岳大・濱本 真一

掲載種別：論文（学術誌）

共著区分：共著

掲載誌名：『理論と方法』34(2)

タイトル：「合理的選択による進学格差の変動モデル：進学率と階層間格差のマクロ・マクロリンク」

発行機関：数理社会学会

発行年月：2020 年 3 月（予定）

著者氏名：濱本 真一

掲載種別：論文（学術誌）

共著区分：単著

3. 外部資金採択実績

採択補助金名：科学研究費

研究種目：若手研究

研究課題番号：18K13096

採択テーマ：「統計・数理モデルによる教育達成格差メカニズムの総合的解明」

研究期間：2018 年 4 月 ～ 2021 年 3 月

資金名称：日本学術振興会

研究代表者名：濱本 真一

西林 勝吾 助教

1. 学会発表

日時：2019 年 6 月 29 日

学会名等：マルサス学会第 29 回大会

場所：沖縄国際大学

発表者：西林 勝吾

発表テーマ：「Fabien Locher (2019) “Neo-Malthusian Environmentalism, World Fisheries Crisis, and the Global Commons, 1950-1970” について（文献紹介）」

2. 研究論文・著書

掲載誌名：Journal of the History of Economic Thought

タイトル：「A.V. Kneese's Water Quality Management Research (1960s), within the History of Environmental Economics」

発行機関：History of Economics Society (HES)

発行年月：2019年9月

著者氏名：西林 勝吾

掲載種別：論文（査読付ジャーナル）

共著区分：単著

3. 外部資金採択実績

採択補助金名：科学研究費

研究種目：若手研究

研究課題番号：00758237

採択テーマ：「A.V.クネーゼを中心とした環境経済学説史研究」

研究期間：2019年4月～2023年3月

資金名称：日本学術振興会

研究代表者名：西林 勝吾

■ 社会調査部会 部会別研究活動等

1. 国際会議出席

日時：2019年5月28日～31日

会議名等：IASSIST Annual Conference 2019

場所：University of New South Wales (in Sydney, Australia)

出席者：五十嵐 彰・高橋 かおり

参加目的：国際会議出席ならびに海外のデータアーカイブ先行機関視察とメタデータ国際標準「DDI」に関する先進動向の調査・聴取および「RUDA」の海外研究者への広報活動

岩間 暁子 教授

1. 研究論文・著書

掲載誌名：『リスク社会の家族変動』（田間泰子編著）

タイトル：「単独世帯の増加と『家族』のオルタナティブ」（印刷中）

発行機関：放送大学教育振興会

発行年月：2020年3月

著者氏名：岩間 暁子

掲載種別：論文

共著区分：単著

*備考）ラジオ教材も別途制作

掲載誌名：『リスク社会の家族変動』（田間泰子編著）

タイトル：「家族の多様化と貧困・社会的排除」（印刷中）

発行機関：放送大学教育振興会

発行年月：2020年3月

著者氏名：岩間 暁子

掲載種別：論文

共著区分：単著

*備考) ラジオ教材も別途制作

掲載誌名：『作業療法ジャーナル』54 巻 4 月号

タイトル：「社会学から見た家族 1」（予定）

発行機関：三輪書店

発行年月：2020 年 3 月

著者氏名：岩間 暁子

掲載種別：論文

共著区分：単著

2. 依頼講演

日時：2019 年 8 月 7 日（10:00－12:00）

講演名：East Asian Program 2019

主催：International Movement of Catholic Students

場所：国立オリンピック記念青少年総合センター

講師：岩間 暁子

講演テーマ：「Child Poverty and Social Exclusion in Japan」

対象：日本,韓国,香港,台湾出身の大学生を中心とした青少年

松本 康 教授

1. 招待講演

日時：2019 年 11 月 1 日

会議名：International Conference: Urban Culture and Spaces of Social Empathy

主催：East Asia Sociological Association & Center for Glocal Culture and Social Empathy

場所：University of Seoul, South Korea

講師：Matsumoto, Yasushi

講演テーマ：Hyper-urbanization in Tokyo: Changing Residential Differentiation and Political Ecology

日時：2019 年 11 月 9 日

会議名：2019 Urban Renewal and Social Governance International Forum

主催：Department of Sociology, Tongji University, City and Society Research Center, Chinese Society of Social Geography Professional Committee, and Fair City and City Governance Summit Team

場所：Shanghai, China

講師：Matsumoto, Yasushi

講演テーマ：Hyper-urbanization in Tokyo: Changing Residential Differentiation and

Political Ecology

2. 書評

タイトル：「吉原直樹著『都市社会学』」『日本都市社会学会年報』37：135—137

発行機関：日本都市社会学会

発行年月：2019年9月

著者氏名：松本 康

掲載種別：書評

共著区分：単著

高木 恒一 教授

1. 研究論文・著書

掲載誌名：『立命館生存学研究』3号

タイトル：「大学アーカイブズの意義と課題：立教大学共生社会研究センターの経験から」発行機関：立命館大学生存学研究所

発行年月：2019年10月

著者氏名：高木 恒一

掲載種別：論文（機関紀要）

共著区分：単著

掲載誌名：『都市政策研究』12号

タイトル：「住宅都市・集合住宅・まちづくり——東京都町田市の事例から」

発行機関：せたがや都市政策研究所

発行年月：2020年3月（予定）

著者氏名：高木 恒一

掲載種別：論文（機関紀要）

共著区分：単著

五十嵐 彰 助教

1. 学会発表

日時：2019年4月7日

学会名等：The 77th Midwest Political Science Association Annual Meeting

場所：Palmer House Hilton in Chicago

発表者：Akira Igarashi, Yoshikuni Ono

発表テーマ：「The Effects of Benefits and Threats of Immigrants on Public Opinion」

日時：2019年4月17日

学会名等：The 4th European Social Survey Conference

場所：University of Mannheim

発表者：Akira Igarashi

発表テーマ：「Compatibility of Ethnic and National Identification Under Multicultural Policies」

日時：2019 年 5 月 26 日

学会名等：移民政策学会 2019 年度年次大会

場所：立教大学

発表者：永吉 希久子・竹ノ下 弘久・五十嵐 彰・木原 盾

発表テーマ：「外国籍者の社会階層に関する全国調査の試み」

日時：2019 年 8 月 29 日

学会名等：The 115th American Political Science Association's Annual Meeting

場所：Washington, DC

発表者：五十嵐 彰・尾野 嘉邦

発表テーマ：「The Effects of Threats and Benefits of Immigrants on Public Opinion」

日時：2019 年 8 月 30 日

学会名等：第 68 回数理社会学会

場所：熊本県立大学

発表者：五十嵐 彰・尾野 嘉邦

発表テーマ：「新自由主義と排外主義：質問紙実験による検討」

日時：2019 年 10 月 5 日

学会名等：第 92 回日本社会学会大会

場所：東京女子大学

発表者：五十嵐 彰

発表テーマ：「移民の心理的統合—在日外国人の社会統合に関する研究（4）」

日時：2019 年 11 月 30 日

学会名等：The 23rd Experimental Social Science Conference

場所：明治学院大学

発表者：Akira Igarashi, Yoshikuni Ono

発表テーマ：「Neoliberalism and Negative Attitudes toward Immigrants」

日時：2019 年 12 月 3 日

学会名等：International Symposium on New and Old Migrations and Diversities in UK and Japan

場所：早稲田大学

発表者：James Laurence, Akira Igarashi

発表テーマ：「Fear and Hoping in Japan: How Immigration Impacts Social and Cultural Life in a Lower-Diversity Society」

2. 研究論文・著書

著書名：『移民政策とは何か：日本の現実から考える』

タイトル：「移民排斥：世論はいかに正当化しているか」(pp.145－pp.165)

発行機関：人文書院

発行年月：2019年4月

著者氏名（共著者含）：高谷 幸（編）・樋口 直人・稲葉 奈々子・奥貫 妃文・榎井 縁・
五十嵐 彰・永吉 希久子・森 千香子・佐藤 成基・小井土 彰宏

掲載種別：著書（書籍）

共著区分：共著

著書名：『日本人は右傾化したのか：データ分析で実像を読み解く』

タイトル：「排外主義：移民増加はその源泉となるか」(pp.94－pp.114)

発行機関：勁草書房

発行年月：2019年9月

著者氏名（共著者含）：田辺 俊介（編）・齋藤 僚介・濱田 国佑・五十嵐 彰・永吉 希久子・米田 幸弘・桑名 裕樹・伊藤 理史・阪口 祐介・松谷 満

掲載種別：著書（書籍）

共著区分：共著

掲載誌名：RIETI Discussion Paper Series

タイトル：「Neoliberalism and Negative Attitudes Toward Immigrants」

発行機関：経済産業研究所

発行年月：2019年10月

著者氏名（共著者含）：五十嵐 彰・尾野 嘉邦

掲載種別：ワーキングペーパー

共著区分：共著

3. 外部資金採択実績

採択補助金名：村田学術振興財団研究助成

採択テーマ：「日本における不倫の規定要因」

研究期間：2019年8月～2020年6月

資金名称：村田学術振興財団

研究代表者名：迫田 さやか

採択補助金名：科学研究費

研究種目：研究活動スタート支援

研究課題番号：19K23254

採択テーマ：「日本におけるエスニックヒエラルキー形成過程の解明」

研究期間：2019年10月～2021年3月

資金名称：日本学術振興会

研究代表者名：五十嵐 彰

高橋 かおり 助教

1. 学会発表

日時：2019 年 8 月 22 日

学会等名：14th Conference of the European Sociological Association

場所：306, Geoffrey Manton, Manchester Metropolitan University

発表者：Kaori Takahashi

発表タイトル：「Building Social Capital in a Port Town Neighborhood with Contemporary Art Projects and Classical Music Concerts」

日時：2019 年 10 月 6 日

学会名：第 92 回日本社会学会（テーマセッション 12「創造性・芸術性と労働をめぐる社会学」）

場所：東京女子大学 6211 教室

発表者：高橋 かおり

発表タイトル：「文化・芸術領域における仕事——遊びと労働の両立可能性の探求」

（フォーラム企画）

日時：2019 年 12 月 21 日

学会名：日本文化政策学会第 13 回研究大会

場所：さいたま市文化センター 展示室

発表者：竹田 恵子・高橋 かおり・山本 浩貴・喜始 宜照

企画タイトル：「アート界におけるジェンダー平等を促進する教育のあり方について」

日時：2020 年 3 月 21 日

学会名：関東社会学会 第 2 回研究例会「ワークショップ時代の統治と社会記述」

場所：明治学院大学白金キャンパス 本館（北ウィング）1254 教室（予定）

発表者：高橋 かおり

発表テーマ：「芸術を通じた場の構築——地域に対する現代美術とクラシック音楽の試みを比較して」

2. 外部資金採択実績

採択補助金名：科学研究費

研究種目：挑戦的研究（萌芽）

研究課題番号：17K18464

採択テーマ：「アート協働制作による社会関係資本形成の社会的・実践的研究」

研究期間：2017 年 6 月～2020 年 3 月

資金名称：日本学術振興会

研究代表者名：高橋 かおり

採択補助金名：科学研究費

研究種目：挑戦的研究（萌芽）

研究課題番号：19K21731

採択テーマ：「グローバル化時代における進学・就職・転職ネットワークの社会学的解明」

研究期間：2019年6月～2022年3月

資金名称：日本学術振興会

研究代表者名：相澤 真一

採択補助金名：2019年度研究助成「日独の若者文化・ライフスタイルの研究」

採択テーマ「ドイツ在住日本人芸術家のキャリア形成に関する比較研究」

研究期間：2019年4月～2020年1月

資金名称：一般財団法人 山岡記念財団

研究代表者名：高橋 かおり

3. 依頼発表

日時 2020年2月19日 13:30～18:00（予定）

講演名：若者文化シンポジウム 第4回研究成果発表会

主催：山岡記念財団

場所：京都大学 第3講義室

登壇者：林 祐一郎・富永 京子・高橋 かおり・シュテファン ヴァルター ブリュクナー・ティモ テーレン

発表テーマ「ベルリン在住日本人アーティストの活動——定住と移動のはざまで」

対象：研究者,一般

■ 統計教育部会 部会別研究活動等

1. 学会発表

日時：2019年9月9日

学会名等：2019年度統計関連学会連合大会

場所：滋賀大学

発表者：山口 誠一・山口 和範・門田 実

発表テーマ：「統計検定を利用した統計教育の改善について」

山口 和範 教授

1. 学会発表

日時：2019年8月21日

学会名等：The 62nd ISI Statistical World Congress

場所：クアラルンプール（マレーシア）国際会議場

発表者：森西 美光・山口 和範

発表テーマ：「Causal Inference in Sport Data ~Causal Model of Fly Ball Revolution in NPB」

日時：2019 年 9 月 18 日

学会名等：International Conference on Japanese Style International Business Program

場所：ハノイ（ベトナム）ハノイ貿易大学

発表者：山口 和範

発表テーマ：「Developing Future Business Leaders for Enhancing Economic Ties Between Vietnam and Japan in the Age of Artificial Intelligence」

2. 外部資金採択実績

採択補助金名：科学研究費

研究種目：基盤研究(C)

研究課題番号：17K00059

採択テーマ：「ビッグデータ時代に求められる統計的思考力育成のための統計教育の研究」

研究期間：2017 年 4 月 ～ 2020 年 3 月

資金名称：日本学術振興会

研究代表者名：山口 和範

3. 依頼講演

日時：2019 年 10 月 16 日 午後 2 時 00 分～4 時 00 分

講演名：日本科学技術連盟 2019 年度 月例講演会

主催：日本科学技術連盟

場所：日本科学技術連盟東高円寺ビル

講師：山口 和範

講演テーマ：「AI 時代における統計の役割と必要とされる統計的考え方」

対象：一般

日時：2019 年 11 月 1 日 午後 1 時 40 分～3 時 50 分

講演名：統計教育講演会

主催：佐賀県

場所：佐賀商業高校体育館

講師：山口 和範

講演テーマ：「統計的に考えるってどういうこと」

対象：1 年生，2 年生，および教員

都築 誉史 教授

1. 学会発表

日時：2019 年 8 月 21 日

学会名等：Subjective Probability, Utility, and Decision Making (SPUDM) 2019 (The European Association for Decision Making)

場所：オランダ，アムステルダム，アムステルダム大学

発表者：都築 誉史・千葉 元気

発表テーマ：「A Time-Series Saccades Analysis of the Attraction and Compromise Effects Based on the Final Decision in Multi-Alternative Decision Making」

日時：2019 年 7 月 19 日

学会名等：Annual Conference of the European Association of Psychology and Law (EAPL) 2019

場所：スペイン，サンティアゴ・デ・コンポステーラ

発表者：島根 大輔・都築 誉史・伊東 裕司

発表テーマ：「Awareness of Emotional Stimuli in a Video: Suppressed Memory of the Video」

2. 研究論文

掲載誌名：Frontiers in Psychology

タイトル：「Effortful Processing Reduces the Attraction Effect in Multi-Alternative Decision Making: An Electrophysiological Study Using a Task-Irrelevant Probe Technique」

発行機関：Frontiers Media SA

発行年月：2019 年 4 月

著者氏名：都築 誉史・武田 裕司・千葉 元気

掲載種別：Original Research Article（査読有り）

共著区分：The First Author

3. 著書

掲載誌名：児童心理学の進歩 2020 年版 (Vol. 59)

タイトル：判断と意思決定

発行機関：金子書房（日本児童研究所・監修）

発行年月：2020 年 6 月（予定，印刷中）

著者氏名：都築 誉史

掲載種別：展望論文（査読有り）

共著区分：単著（分担執筆）

山口 誠一 助教

1. 学会発表

日時：2020 年 3 月 12 日－14 日

学会名等：32nd International Conference on Technology in Collegiate Mathematics

場所：Orlando, FL, USA

発表者：Satoru Takagi, Kesayoshi Hadano, Sei-ichi Yamaguchi

発表テーマ：「Educational Materials on Calculus Which Bridge Mathematics and

Engineering]

2. 依頼研修

日時：2019 年 5 月 8 日, 15 日, 22 日, 29 日, 6 月 5 日, 12 日 18 時 15 分～19 時 45 分, 19 時 55 分～21 時 25 分

講演名：統計検定 2 級資格取得講座

主催：東洋大学経営学部

場所：東洋大学白山キャンパス 3301 教室

講師：山口 誠一

講演テーマ：第 1 回～第 6 回

対象：JINSE 版統計検定に受験申込みを行った東洋大学の学生

日時：2019 年 10 月 3 日, 10 日, 24 日 午後 6 時 10 分～8 時 10 分

講演名：統計検定 2 級試験対策講座（第 2 回, 第 3 回, 第 5 回）

主催：東洋大学理工学部

場所：東洋大学川越キャンパス 521 教室

講師：山口 誠一

講演テーマ：「記述統計（第 2 回）」, 「確率変数（第 3 回）」, 「推測統計（推定）（第 5 回）」

対象：JINSE 版統計検定に受験申込みを行った東洋大学理工学部の学生

日時：2019 年 10 月 8 日, 15 日, 22 日, 29 日, 11 月 12 日, 19 日 18 時 15 分～19 時 45 分, 19 時 55 分～21 時 25 分

講演名：統計検定対策講座

主催：東洋大学経営学部

場所：東洋大学白山キャンパス 3404 教室

講師：山口 誠一

講演テーマ：第 1 回～第 6 回

対象：JINSE 版統計検定に受験申込みを行った東洋大学の学生

門田 実 助教

1. 依頼研修

日時：2019 年 8 月 27 日～2019 年 8 月 29 日

研修名：データ分析研修

主催：北海道大学海洋生物資源学部海洋資源科学科

場所：北海道大学

講師：門田 実

対象：北海道大学海洋生物資源学部海洋資源科学科

2. 依頼講演

日時：2019 年 11 月 21 日 午後 6 時 00 分～7 時 30 分

講演名：サルでもわかる最先端データ分析のミライ学

主催：四万十町役場人材育成推進センター

場所：古民家カフェ「半平」

講師：門田 実

講演テーマ：「2016年度の経営環境と情勢に負けない企業づくり」

対象：AIに興味のある方,データ分析に興味のある方

加藤 倫子 教育研究コーディネーター

1. 外部資金採択実績

採択補助金名：科学研究費

研究種目：若手研究

研究課題番号：19K13951

採択テーマ：「構造的排除への支援・介入が「犯罪や非行からの離脱」に与える影響」

研究期間：2019年4月～2023年3月

資金名称：日本学術振興会

研究代表者名：加藤 倫子

2. 依頼研修

日時：①2019年5月8日 午後1時10分～午後3時

②2019年5月15日 午後1時10分～午後3時

研修名：個人探求・RQを考えよう

主催：豊島岡女子学園高等学校

場所：豊島岡女子学園高等学校

講師：加藤 倫子

研修テーマ：『『犯罪から立ち直る』とはどのようなことか』

対象：豊島岡女子学園高等学校2年生

社会情報教育研究センター研究紀要規程

2014年12月1日制定

2017年1月16日改正

2018年1月17日改正

I. 名称

1. 本誌「社会情報教育研究センター研究紀要『社会と統計』」は立教大学社会情報教育研究センターの機関誌であり、原則として年1回発行を行うものとする。

II. 目的

2. 本誌は社会情報教育研究センターにおける研究教育成果を公表するとともに、統計・社会調査・ICTを活用した研究教育支援の高度化に寄与することを目的とする。

III. 内容

3. 本誌は社会情報教育研究センターの活動記録（講演会・シンポジウム・研究会等を含む）、研究論文、研究ノート、調査報告、資料紹介、書評、その他編集委員会が認めたものを掲載する。

IV. 編集委員会

4. 本誌編集は「社会情報教育研究センター研究紀要『社会と統計』」編集委員会が行う。
5. 紀要編集にかかる編集委員については、センター長が任命する。
6. 編集委員会は、原則として社会情報教育研究センター政府統計部会・社会調査部会・統計教育部会の各リーダー、または各リーダーが推薦する者およびCSI事務局を構成員とし、その活動にあたるものとする。

V. 投稿資格

7. 投稿資格は以下の通りとする。
 - (ア) 資格区分Ⅰ 当センターに所属する教職員等（任期付も含む）
 - (イ) 資格区分Ⅱ 当センターにおいて研究・教育に従事している、又はしたことのある者
 - (ウ) 資格区分Ⅲ 本学大学院研究科博士課程後期課程在籍者
 - (エ) 資格区分Ⅳ その他、編集委員会が認めた者

VI. 掲載基準

8. 編集委員会の判断に基づき、掲載の採否を決定する。

VII. 原稿の提出

9. 原稿は編集委員会が設定した期限に合わせて、提出を行う。

- 1 0. 校正は著者校正とする。提出された原稿等は返却しない。
- 1 1. 原稿は日本語あるいは英語で執筆された未刊行のものとするが、研究会等で口頭発表したものについてはその限りではない。
- 1 2. 詳細は別途定める。
- 1 3. 他の文献から図・表・写真の転載を行う場合は、執筆者が許諾を必ず得ること。

VII. 著作権

- 1 4. 本誌に掲載された原稿のすべての著作権（著作権法第 27 条及び第 28 条に規定する権利を含む）は当センターに譲渡される。
- 1 5. 本誌に掲載された原稿は、原則として立教大学学術リポジトリ（立教 ROOTS）を通じてオンライン公開を行う。
- 1 6. 立教大学学術リポジトリ（立教 ROOTS）での公開を希望しない場合は、原稿提出時に編集委員会にその旨を通知すること。

IX. その他

- 1 7. その他本誌の編集に関して、必要な事項は編集委員会で審議を行う。
- 1 8. この規程の改廃は、センター委員会の議を経て、センター長が行う。

社会情報教育研究センター研究紀要執筆要領

2014年12月1日制定
2017年1月16日改正
2017年10月4日改正
2019年1月16日改正

I. 原稿の形式

1. 原稿は1頁の字数を1600文字(40×40)とし、概ね以下の通りとする。
作成は指定された執筆用テンプレートを使用し、執筆すること。
 - ① 論文は15頁程度
 - ② 研究ノートは10頁程度
 - ③ 資料は10頁程度
 - ④ その他掲載については、編集委員会の許諾を得て、掲載方法等を決定する。なお、原稿に図表が含まれる場合には、紙面に占める割合を加味し、構成する。
2. 使用言語は原則として日本語もしくは英語とする。
3. 原稿の冒頭に、表題(英文・和文)と執筆者の氏名を記載する。副題(英文・和文)があれば、表題の下部に記す。
4. 和文要旨は概ね400字を目安とする。アブストラクト(英文)は概ね200wordを目安とする。
5. キーワード(英文・和文)は各5個以内で設定する。
6. 節や項等の階層構造による見出しは、以下の形式を基準とする。ただし、必要に応じて構成は柔軟に対応できるものとする。
大項目「章」相当：Ⅰ、Ⅱ、Ⅲ・・・
中項目「節」相当：1、2、3・・・
小項目「項」相当：(1)、(2)、(3)・・・(数字はいずれも半角とする)
7. 図表や写真などは「図表1、図表2・・・」というように通し番号をつける。(数字は半角とする)
8. 「注」については、本文中の該当箇所に通し番号で1)2)・・・と付け、注自体は本文の最後にまとめて掲載する。
9. 文中で使用する引用文献は、本文の該当箇所に(著者名、発行年)を示し、詳細は論文末尾に一括して記載する。
10. 引用文献掲載形式は概ね以下の通り。
 - (1) 著書：著者名、出版年、『書籍名』出版社.
 - (2) 論文：著者名、出版年、「論文名」『ジャーナル名』○巻○号、1・2(ページ).
 - (3) ウェブページ：著者名、最終更新年、「タイトル」、ウェブサイト名、(URL, ○年○月○日取得)
11. 論文等は、完全原稿で提出し、その内容については執筆者本人が責任を負うものとする。

II. その他

12. 本要領は、社会情報教育研究センター研究紀要規程12に基づく。
13. 本要領の改廃は、社会情報教育研究センター編集委員会の議を経て、センター長が行う。

執筆者（掲載順）

- ◆ **松本 康（まつもと やすし）**
立教大学 社会学部 教授
- ◆ **菊地 進（きくち すずむ）**
立教大学 名誉教授
- ◆ **山口 和範（やまぐち かずのり）**
立教大学 社会情報教育研究センター 統計教育部会長
立教大学 経営学部 教授
- ◆ **水上 徹男（みづかみ てつお）**
立教大学 社会情報教育研究センター センター長
立教大学 社会学部 教授
- ◆ **櫻本 健（さくらもと たけし）**
立教大学 社会情報教育研究センター 政府統計部会長
立教大学 経済学部 准教授
- ◆ **濱本 真一（はまもと しんいち）**
立教大学 社会情報教育研究センター 政府統計部会 助教
- ◆ **西林 勝吾（にしばやし しょうご）**
立教大学 社会情報教育研究センター 政府統計部会 助教
- ◆ **山口 誠一（やまぐち せいいち）**
立教大学 社会情報教育研究センター 統計教育部会 助教
- ◆ **門田 実（かどた みのもる）**
立教大学 社会情報教育研究センター 統計教育部会 助教
- ◆ **五十嵐 彰（いがらし あきら）**
立教大学 社会情報教育研究センター 社会調査部会 助教
- ◆ **高橋 かおり（たかはし かおり）**
立教大学 社会情報教育研究センター 社会調査部会 助教

紀要編集・構成

- ◆ **加藤 倫子（かとう みちこ）**
立教大学 社会情報教育研究センター 教育研究コーディネーター
- ◆ **小山田 基香（おやまだ もとか）**
立教大学 社会情報教育研究センター 教育研究コーディネーター

編集委員会

編集委員長 水上 徹男 (社会情報教育研究センター長 社会学部 教授)

政府統計部会 編集委員 櫻本 健 (経済学部 准教授)

社会調査部会 編集委員 岩間 暁子 (社会学部 教授)

統計教育部会 編集委員 山口 和範 (経営学部 教授)

社会情報教育研究センター 事務局 編集委員

毛利 立夫 (メディアセンター 課長 社会情報教育研究センター事務局 兼務)

重田 根見子 (メディアセンター 課員 社会情報教育研究センター事務局 兼務)

木田 英樹 (メディアセンター 課員 社会情報教育研究センター事務局 兼務)

宮本 信愛 (メディアセンター 課員 社会情報教育研究センター事務局 兼務)

加藤 倫子 (社会情報教育研究センター事務局 教育研究コーディネーター)

小山田 基香 (社会情報教育研究センター事務局 教育研究コーディネーター)

立教大学社会情報教育研究センター研究紀要

社会と統計 第6号

2020年2月28日 発行

編集・発行：立教大学 社会情報教育研究センター

発行責任者：水上 徹男

〒171-8501 東京都豊島区西池袋 3-34-1

TEL : 03-3985-4459 FAX : 03-3985-2907

Email : csi-info@rikkyo.ac.jp URL : <https://spirit.rikkyo.ac.jp/csi>

印刷：株式会社 三恵社

〒462-0056 愛知県名古屋市中丸町 2-24-1

TEL : 052-915-5211 FAX : 052-915-5019
