社会調査の公開データ ----2次分析への招待

佐藤博樹・石田 浩・池田謙一編 東京大学出版会 2000年 A 5 判 260頁 2,800円

たとえば NHK 放送文化研究所編『現代日本人の意識構造』(日本放送出版協会)は、5年ごとに実施される大規模なアンケート調査をもとに「今の日本人のものの見方はどのようなものか、そしてそれは過去と比べてどのように変わった(変わっていない)のか」を多面的に把握しようとする企てであり、我が国における代表的な社会調査のひとつといえる。

アンケートに答えるのは個々人だが、基本的に「どの個人がどんな回答をしたか」までは特定されない。性別、年齢などといった条件によって回答の傾向にどのような差異が生じるかは分析の対象となるが、税務調査や学校での試験とは異なり、個人を特定する固有名詞そのものが必須データとなることはない。

調査結果の信頼性を確保するにはかなり大規模なデータ収集が必要であり、一個人が気やすく行なえるものではない。「××の問題について、一般的傾向として日本人はどのように考えているか」を知ろうと思えば、性別、年齢、職業や居住地域など、いくつかの属性においてバランスのとれたデータ(個票)を十分に集めなければならない。大きな課題に取り組もうとすればするほど、分析の信頼性を確保するために要する労力は大きくなっていく。しかるべき組織や機関によって信頼性の高いデータが蓄積・公開されているならば、それに依拠することで調査労力を省き、データ分析の方にエネルギーを集中できる。これが社会調査データ公開の第一の意義だろう。

また多くのばあい調査報告は分析結果を提供するにとどまり、もとデータそのものまで公表されることは少なく、分析の妥当性を検証することがむずかしい。加えて同じデータ

を別の観点から分析したいと思っても、通常 は収集作業にあたった当事者でなければ、 データを入手することは困難なことが多かっ た。

つまり、後日の検証可能性を担保すること、 第三者による別の観点からの分析の機会を確 保することもまた、調査データ公開の重要な 意義であろう。そして一方で、近年における コンピュータおよびソフトウェアの飛躍的な 発展は、データの集計・比較といった「作業」 労力を画期的に軽減しつつある。

これらを背景に本書は公刊された。副題にある「2次分析(secondary analysis)」は本書冒頭では「最初の研究(1次分析)では明らかにされなかった点を解明する研究」と説明されているが、すでにアメリカなどでは調査(データ収集)を終了して短期間のうちに、集計・分析などの操作を経ることなく早々に個票データが公開されるようになっている場合もあり、そうなるともはや1次、2次の境界は曖昧となる。「研究の素材として公開・提供されている(自力で集めたわけではない)調査データを利用して行なわれた分析」と理解しておけば大過ないだろう。

本書の構成は以下のとおり。

序

- I 公開データによる研究・教育の方法
 - 1 公開データによる社会分析の手引き
 - 2 公開データ利用型の調査教育の勧め
- II 公開データセットの紹介
- 1 データ・アーカイブとは何か
 - 2 公開データの概要
- 3 主要な公開データ
- III 公開データを利用した社会分析
 - 1 社会的資源と地位達成
 - 一アメリカの総合社会調査による分析
 - 2 政治的疎外意識と政治参加
- 一「衆議院選挙に関する世論調査」によ る分析
 - 3 旧東西ドイツにおける職業生活価値 観の違い
 - 一ドイツの総合社会調査による分析

- 4 出産・夫の育児と妻の夫婦関係満足 度
 - ─「消費生活に関するパネル調査 (JPSC) | による分析
- 5 働く母親と子どもの親子関係
 - -ISSP 調査による分析
- 6 福祉国家は今なお支持されているか —ISSP調査による分析
- 7 結婚観の国際比較 —ISSP 調査による分析
- IV 公開データ入手の方法
 - 1 公開データの入手方法
 - 2 資料 関連機関のホームページアドレス

あとがき

I 部では数量化されたデータをもとにした 実証研究の手法およびその教授法に関する見 解が述べられる。本書の立場から重要なのは 後者だろう。大学などでの社会調査教育がこ れまで「収集過程重視型教育」に偏し、えて して調査は大変なばかりで面白くないという 印象を学生に残しがちであったという指摘に 続けて、データ分析の面白さを実感させ、そ の上でデータ作成・処理のプロセスを教授す る「公開データ利用型教育」へとシフトする ことの意義が強調されている。

II 部では国内と海外(主として欧米) について、データ・アーカイブ整備状況の概観と、 具体的にどのような領域のデータが公開されているかに関する紹介が行なわれている。

まず冒頭で、データ・アーカイブとは「社会調査のマイクロデータ(個々の調査票の記入内容のデータ)を収集・保管し、学術目的での2次分析のために提供する機関である」と定義され、各国の現状がまとめられている。

アメリカでは20世紀半ばに民間の世論調査会社が商業ベースのデータ・アーカイブを設立し、その後1960年代には複数の大学・研究機関が構成するコンソーシアムによって別のデータ・アーカイブも設立され、参加大学の研究者、学生が自由にマイクロデータを利用

できるようになった。前者はヨーロッパでのデータ収集活動にも乗り出し、これも一因となってヨーロッパ独自のデータ・アーカイブが設立されることとなり、1960年代はじめに西ドイツ(当時)、イギリス、オランダでデータ・アーカイブの活動が開始された。以上アメリカの2機関、ヨーロッパの3機関が社会科学分野における欧米の主要なデータ・アーカイブに数えられている。他のヨーロッパ諸国でも設立が進められる一方、ヨーロッパ各国のデータ・アーカイブ間の協力をはかるための組織も1976年に発足する。

日本でも社会調査そのものは熱意ある研究者によって進められてきたが、そのもとデータの保存・公開という面では欧米に後れをとっていた。2次分析に供することを前提とした社会調査マイクロデータの公開が大学レベルで進められるようになったのは、1990年代に入ってからのことである。

III部は題目どおり現在入手可能な公開データを利用してなされた諸研究の紹介である。 煩を承知で個別タイトルの紹介に及んだのは、現在すでに公開されているものを利用するだけでも、ずいぶん多様な領域について調査データを得ることができるようになっているということを示そうとしたためであった。

まず日本、アメリカ、イギリス、ドイツそれぞれにおける総合社会調査が紹介される。総合社会調査(GSS: General Social Survey)とは、社会学的、心理学的観点から社会の趨勢を把握するため、「経済指標」に匹敵するような「社会指標」を構築し、それを公共政策に反映させ、社会科学の知識を有効に活用しようとする社会指標運動を発端として1970年代初頭のアメリカで開始され、その後ヨーロッパや日本でも始められたもので、さまざまな問題関心からの分析が可能となるような調査データの構築がめざされている。

つづけて価値観やジェンダー、選挙や消費 生活などに関して実施された調査で、データ 公開されているものをもとにした社会分析が それぞれ試みられている。 IV部ではどのような社会調査データが公開されているか、どのような方法でそれを入手できるかに関するガイドが行なわれ、関連機関のホームページアドレスが紹介されている。

こう見てくると、社会調査データの公開といっても、利用者として想定されているのがもっぱら研究者や学生だという点で、一般の市民を利用者として想定する文書館とは確かに異なる。文書館のばあい資料にアクセスできるかどうかが当該資料の性格に左右されることはあっても、館の利用そのものが希望者の属性によって制限されるとは考えにくい。

実際の例をドイツの ZA (Zentralarchiv) に見てみよう。この機関では、データの利用 範囲を次のように分類しているという。

- O 誰もが利用可
- A 高等教育機関、公的機関等での教育ま たは研究に利用可
- B 結果を発表しなければ高等教育機関で の教育または研究に利用可。結果を発表 する場合またはその他の機関で利用する 場合には寄託者の許可が必要
- C 寄託者の許可の下で教育または研究に 利用可

誰の利用か、なんのための利用かに応じてかなり厳密な制限の設定されていることがわかる。「公開データ」とはいえ、それだけ取り扱いには注意が必要なのだといえるだろう。

公開の対象となる調査データは、たとえ個票に氏名が明記されていなくても、調査時期、居住地域、年齢など周辺情報によって個人を特定することも可能なものであり、そうである以上、その不注意な扱いはプライバシー広くは人権の侵害にもつながりかねない。情報公開とプライバシー保護の関係をどうバランスさせるかという課題は文書館にも共通と考えられるが、ここでは、教育・研究に利用を限定し、さらに結果の公開に寄託者の許可が必要というかたちでそれを担保しようとしている。

この意味で、政府はじめ行政機関が主体となって実施する調査データの処遇をどうする

かは重要な問題となりそうだ。調査、分析には税金が投入され、また調査の包括性という点からも、そのデータは国民、住民の貴重な共有財産といえる。しかし国勢調査を想起すれば端的に了解されるように、政府統計の多くはあらかじめ設定された目的以外での利用が法律によって厳しく制限されている。

すでに公開されているデータの入手・活用 については本書から周到な議論と案内を得ら れる。どのようなデータをどのようなかたち で「公開」したらよいかについては、今後なお 慎重かつ精力的な議論が必要となるだろう。

最後に、個票データ2次利用の効用という点で、佐藤俊樹『不平等社会日本一さよなら総中流』(中公新書)に触れておきたい。これは、戦後から高度成長期を経て50年たった現在の日本が「努力すればナントカなる」=「開かれた社会」から「努力してもしかたない」=「閉じた社会」に変容しようとしていることを示し(ときに斎藤貴男『機会不平等』(文藝春秋)とも並べられつつ)話題となった書物であるが、まさに本稿で紹介してきた2次分析という作業の成果だともいえる。

1955年以来10年おきに、全国の人々を対象として当人および両親の職業、学歴、社会的地位など、社会階層にかかわるさまざまなデータを集めてきている「社会階層と社会移動全国調査」通称 SSM (Social Stratification and Mobility) 調査の1995年研究メンバーとして加わった著者は、過去の個票データを再整理するという作業にも取り組み、その中からこのような成果を生みだした。進行中の調査に携わっていた著者だからこそ、過去のデータを眺める視点も豊かなものになったという部分があるだろうと推測される。

このような成果が生みだされるのであれば、利用者、利用目的にあるていどの制限を 設けたうえでもなお、調査データを公開する ことの意義が揺らぐことはないだろう。

なお、「アーカイブ」「アーカイブズ」それ ぞれの語を「収集されたデータ(資料)」「デー タを集積・提供する機関」のいずれに当てる

かには議論もあるが、ここでは本書にならっ た。「人々の共有財産を保管・整理して広く利 用、鑑賞に供する機関 | という意味では文書 館の他にも図書館、博物館、美術館などを想 起することができる。これに社会調査のデー タ・アーカイブ (調査データの保管・提供機 関) も含めたうえで、それらの関係がどのよ うに位置づけられるのかについて、残念なが ら筆者自身はよく知らない。行政の一機関と しての文書館であっても、それが「文書館」 「公文書館」いずれの呼称を与えられるかに よって、人々の持つイメージはおのずと変 わってくるはずだから、「共有財産を集積・提 供する機関」という枠のなかでの個々の関連 性の位置づけが将来的には必要になってくる だろうと推測される。

もちろん、同じ「人々の共有財産」といっても、それが近代以前の歴史的文書であるか、近代の行政文書であるか、あるいは美術品、民俗資料、出版された書籍などのいずれであるかによって、保管・整理の方法も提供・公開の方法も一様ではないし、また対象とする「人々」の範囲をどう設定するかという。も機関によって異なっている。現状では、それぞれの扱う資料の性格に沿って特化した方向で施設の整備が行なわれてきていると見ていだろう。ここで紹介した社会調査のデータ・アーカイブなどは「研究者の研究者による研究者のための(一般市民の利用を前提としない)アーカイブ」と位置づけることも、あながち無謀とはいえないように思われる。

共有財産の公開といったばあい時代の流れからしばしば「インターネット上での公開」が視野に入れられる。ここに紹介した社会調査データのアーカイブも、その多くがインターネット経由でのデータセットの入手を前提としている。こうしたことが契機ともなって所蔵資料管理・公開のための方法をできるだけ汎用性の高いものにしようという動きも出てくるだろう。そのような動きのなかで、異なる対象を扱う機関の管理・公開の方法が標準化されていくのだろうか、反対に方法の

個別化・特殊化が進むのだろうか。これは、それぞれの機関で求められるスタッフの専門性をどのようなものと位置づけるか、そして、そうしたスタッフ確保のために、どのような供給源が用意されるかということにも左右されるだろう。興味深い点ではあるが、これ以上は書評の範囲を超える問題だ。

藤吉圭二・高野山大学/社会学