

■ 研究ノート ■

社会調査データのアーカイブ化をめぐる現状と課題について¹⁾

福内千絵*

1. はじめに

関西学院大学先端社会研究所では、「文化的多様性の尊重する社会の構築」という基本テーマのもと、これまで数多くの研究プロジェクトが立ち上げられ、社会調査を基軸とした研究活動が進められてきた。個々の研究プロジェクトにおいて社会調査データが収集され、それに基づいた研究成果は、論文や著書および叢書として公刊されている²⁾。本研究所設立から10年が経過しようとしている現在、こうして収集・蓄積された社会調査データはどのように「保存、そして公開・共有」すればよいのか、すなわちそのアーカイブ化の方向性について、実務的・理論的観点から検討すべき段階に入っている。本稿では、今日の社会学分野におけるアーカイブ化をめぐる現状と議論を整理することで、社会調査データのアーカイブ化における課題について考察する。

2. 社会調査データのアーカイブ化をめぐる現状：量的調査と質的調査

日本ではこれまで、さまざま地域や現場での調査のほか、国際的な比較調査や大規模な量的調査など、数多くの社会調査が行われ、大量の調査データが蓄積されてきた。そして、学術資源としての調査データの保存・公開に対して意識が向けられるようになり、アーカイブ・ルールの確立やシステムの設立についての議論もおこなわれるようになった。また近年、インターネット環境が進歩したことによって、量的調査を中心としたデータの公開性が飛躍的に高まり、データアーカイブによって公開された調査データの二次利用や共有化が進んでいる。日本における量的調査データのアーカイブに関しては、主導的・代表的なものとして、東京大学社会科学研究所附属社会調査・データアーカイブ研究センターが運営するSSJデータアーカイブ(SSJDA)があり、公開された統計

*関西学院大学先端社会研究所 専任研究員

- 1) 筆者は、先端社会研究所のデータアーカイブ研究部門における研究活動の一環として、人間文化研究機構国文学研究資料館主催による「平成29年度アーカイブズ・カレッジ(史料管理学研修会通算第63回)短期コース」に参加し、所定の課程を修了した(修了証書第28号 平成30年1月31日)。本稿はカレッジ修了論文「社会調査データのアーカイブ化をめぐる現状と課題について」に若干の加筆修正を施したものである。
- 2) 成果公刊の一例として、2012-2015年度共同研究における叢書『「排除と包摂」を超える社会理論』(全3巻、明石書店、2017年)がある([第1巻]:『中国雲南省少数民族から見える多元的世界——国家のはざまを生きる民』(荻野昌弘・李永祥編)、[第2巻]:『在日コリアンの離散と生の諸相——表象とアイデンティティの間隙を縫って』(山泰幸編)、[第3巻]:『南アジア系社会の周辺化された人々——下からの創発的生活実践』(関根康正・鈴木晋介編))。

データや社会調査の個票データを用いた多様な二次分析の成果が産出されている³⁾。

一方、質的調査データについては、量的データに比べて、より多様な形態のドキュメントをふくむことから、データの整理・管理に膨大な作業を必要とする。そのため、日本国内では質的調査のデータアーカイブの設置や質的データ二次分析は、量的データ二次分析ほど進んでいないのが現状であるとされる（桜井 2014；森本 2014；武田 2015）。また、質的調査データアーカイブの設置が進まない理由として、質的調査では、研究者個人によってデータが収集・保管されるため、研究者が亡くなるとそのデータは所在不明となることが多いことや、フィールドワークやインタビュー調査を中心とする質的調査の記録データ（フィールドノート、録音、映像など）には、個人情報が多く含まれており公開になじまないと考えられるためであることも指摘されている（森本 2014：125）。こうした現状のなかで近年、日本の社会学分野では、蓄積された調査資料やデータに関して、「リサーチ・ヘリテージ（調査遺産）」という概念を導入したり、アーカイブ化が進んでいる海外の動向を参照することで、データ・アーカイブの方向性を探る動きがみられる。武田尚子によると、質的調査データのアーカイブ化について、とくに参考になるのがイギリスで、質的データ二次分析に関する先進的な取り組みが行われているという（武田 2015：71）。イギリスでは国立機関の中に、質的データの「収蔵・公開・二次利用・二次分析」を促進する専門部局・ESDS-Qualidataが設置されている。質的データ二次分析の研究環境の構築が積極的に進められてきた背景には、イギリスにおけるワーキング・クラス研究、オーラル・ヒストリー研究の貴重なデータ蓄積が散逸してしまうことへの懸念があったとされる（武田 2015：71）。

以上見てきたように、社会調査データの公開・共有化・二次利用については、すでに量的調査において進められてきており、質的調査においても、データの保存や二次分析の研究環境の構築をはじめとした同様の取り組みが模索されつつあるというのが現状である。さらに次項では、少し視野を広げ、学術全般の研究データの取扱いをめぐる近年の動向に目を向けることで、社会調査データのアーカイブ化の現状を位置付けてみたい。

3. 研究データの取扱いをめぐる動向：公開性、Data Management Plan の重要性

研究データの取扱いをめぐる近年の動向については、前田幸男の報告⁴⁾が非常に参考になる。以下の記述は前田氏の報告に基づいてまとめたものである。

国際機関 OECD（経済協力開発機構）は、公的資金によって収集されたデータは公開することを勧めている。日本ではそれに呼応した動きが見られず、文部科学省や学術会議がデータを二次利

3) 東京大学社会科学研究所附属社会調査・データアーカイブ研究センターは、前身の日本社会研究情報センターから 2009 年 4 月に組織変更し、2010 年 4 月から文部科学省が認定する社会調査・データアーカイブ共同利用・共同研究拠点として活動している。本センターの活動に関連する論文に、佐藤博樹「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』がある。

4) 前田幸男「社会調査個票デジタルデータの収集、保存、二次分析について」14th International Conference on Digital Preservation. Kyoto, Japan, September 25-29, 2017, 報告資料。ipres 2017.jp/wp-content/uploads/jtutorial_maeda.pdf (2017 年 12 月 11 日アクセス)

用のために公開することを義務付けてはいない。研究データの公開に関して、文部科学省は「研究活動における不正行為への対応等に関するガイドライン」を公表（2014年8月）したが、これは不正防止を念頭に置いており、研究データの共有・再分析を推奨するものではない。現在、日本の社会科学系の学会でデータの公開を義務付けたり、論文投稿時にデータの提出を求めている例はないと思われる。研究データの取扱いに関して、諸外国では資金提供機関が Data Management Plan（以下 DMP）を要求することが多い。すなわち資金提供を受ける研究者は、研究計画の DMP の中にデータの保存や公開について記述することが求められるのである。こうしたなかで日本でも、科学技術振興機構（以下 JST）が 2017 年 4 月に「『オープンサイエンス促進に向けた研究成果の取扱いに関する JST の基本方針』運用ガイドライン」において DMP を義務づけた。JST によれば DMP は、「研究者が自身で実施する研究プロジェクトにおいて研究成果として生じる研究データの取扱いを定めるものであり、具体的には、当該研究データの保存・管理、公開・非公開、公開範囲等に関する方針や計画について記載したものを指す」とされる。日本学術振興会も現在、DMP の導入を検討中のようなのである。

研究者にとっては、データの保存・公開のための作業は直接研究に結びつかない「無駄に見える」作業と考えられるが、これからはデータが公開されることを前提に、必要な経費や人件費を研究計画の中に組み込み、計画およびデータ収集の各段階で必要なメタデータを整理・保存するほうが格段に効率的であることも前田によって示唆されている。

以上前田幸男の報告に基づいて研究データの取扱いをめぐる動向をみてきた。「公的資金から得られたデータは公開する」という方針は、次項で述べる「公共性」の観点からの社会調査データのアーカイブ化に関する議論にも接合するものである。また、世界的にはより多くの研究データが保存され、広く二次分析に供されるために DMP が重要視されており、日本でも様々な研究分野において DMP 導入の動きが活発化することが予想される。社会科学の分野においても、DMP に基づいた社会調査データの公開が要請されると考えられ、DMP の内容（公開／非公開・公開範囲など）についてアーカイブ化と連動した議論が求められるところである。このように日本においても研究データの取扱い方針の明文化や研究データの公開が勧められる状況にあるなかで、社会調査データを公開・共有化することにどのような意義や問題点が考えうるのかについて、次項でみていく。

4. 社会調査データの公開・共有化の意義と問題点

社会調査データのアーカイブ化の主たる目的は、二次利用・二次分析にあるとされ、学術的観点からその意義が議論されてきた。第2項でみたように、量的調査データ・アーカイブについては、すでに二次分析による成果報告が進んでおり、その学術的価値は揺るぎないものとなっている。一方、質的調査データに関しては、データの形態が多様であることから、その公開や共有化の学術的意義については慎重に議論されている。すなわち質的調査データは、1. インタビュー、2. 映像資料（写真、動画）、3. 文字資料、4. 自身のフィールドノート、5. その他の資料、という5種類に分けられるが、そのうちインタビュー・データに関しては、「その特性を踏まえるとインタビュー・コンテキストから切り離されたデータの価値はいかに確保できるのか」（小林 2014: 120）と

いった論点が浮上すると考えられる。また、こうした多様な形式による質的調査データについて、その二次利用・二次分析には、おもに次の6つの方法があるといわれている。「第一はオリジナル調査データを歴史的・一次史料として用いる方法、第二は比較研究に素材として活用する方法、第三はオリジナル調査の時期には登場していなかった新しい概念や視点で、調査データを再解釈する方法、第四はオリジナル調査の設計・調査方法の再評価、第五はオリジナル調査の分析内容の妥当性の検証、第六はオリジナル調査データを教材として活用する方法である」(小林2014:115)。

こうした様々な二次利用・二次分析の方法をもつ質的調査データの公開について、社会学者の森本一彦は、フィールドワーカーが収集したデータを、時間を経て歴史社会学者を含む歴史研究者が利用する可能性を示唆し、研究資料の存在を研究領域に関係なく多くの研究者に公開することこそが重要であるとして、学際的な資料価値の観点から質的調査データ公開の意義を述べている(森本2014:131)。

上述した点を踏まえて、質的調査データのアーカイブ化の実務レベルにおいては、多様な研究領域での利用や国際的な学術研究への貢献を視野に入れ、目録データ仕様を考える必要があるが、こうしたアーカイブの実装については別稿において改めて考えたい⁵⁾。

また「公共性の高いデータ」という観点で、調査データの種類によってデータの公開性を考えるべきという主張もあり、社会調査データの公開・共有化の意義はその公共性の高さとも関連して確認されるといえる。調査データならなんでも同様にとらえるのではなく、調査データの内容や調査の資金源に従って高い公共性をもつデータとそうではないデータがある。たとえば調査が公的資金によるものや広く関心を集めそうな調査内容である場合、それらの調査データは公共性が高いと考えられるべきであり、そのようなデータはアクセス可能性が担保され、他者による分析へ開かれるべきであるという(小林2014:115)。この点は第3項でふれた研究データの公開をめぐる世界的動向とも結びつく考え方である。ただし、さらに研究データを公開する時期については議論すべき点があることを指摘したい。「ある研究者(集団)が資金と人的労力を投資して得たデータを第三者が見ることができることになるが、これでよいのか⁶⁾」という問いである。初めからデータを公開すべきか、それともデータの制作に関与した人がまずは分析、論文出版をしてからがよいのか、といった公開の時期的段階については自然科学系の分野のみならず学術全般において今後の論点となると考えられる。

このように質的調査データのアーカイブ化については公開に前向きな意見がある一方で、公開の

5) メタデータの記述に関しては、古賀崇講師の講義「アーカイブズと情報コントロール」(「平成29年度アーカイブズ・カレッジ」2017年11月15日)やその際に示された参考文献において多くの示唆を得た。社会調査データのメタデータ作成については国際標準 ISAD (G): General International Standard Archival Description に基づき、またウェブ上の資料・情報を記述するためのメタデータの枠組みとしては「相互運用性」の高い「ダブリン・コア Dublin Core」の採用が望ましいと考えられる。また、学術研究活動におけるデジタル・アーカイブの影響や意義については、石田佐恵子、山田富秋らによる「特集『映像アーカイブズを利用した質的調査の探求』によせて」(『社会学評論』65巻4号、2014年、pp.454-464、DOI: <https://doi.org/10.4057/jsr.65.454>)において、相互参照に資するデジタル・アーカイブの利点についても言及されており、とくに質的社会調査のデータ・アーカイブの実装を考える上で参考になる。

6) 菊谷英司講師の講義「自然科学系のアーカイブズ」(「平成29年度アーカイブズ・カレッジ」2017年11月6日)において指摘された。

意義や必要性を十分認識しながらも、プライバシーへの配慮から、また調査協力者との固有の関係性のゆえにデータの保存管理は調査者個人によるほかないという意見もある（桜井 2014 : 112）。「公開性、公共性」と「プライバシーや個人情報保護の配慮による秘匿性と個人保管」という矛盾を、質的データが抱えているということはデータ・アーカイブ化の方針を定める上で留意すべき重要な点である。

5. むすびにかえて——アーカイブ化における課題

以上みてきた質的データが抱える「公開性、公共性とプライバシーや個人情報保護の配慮による秘匿性と個人保管という矛盾」を克服する方途として、小林多寿子は、質的調査データの公共性の認識のもとに「アーカイブ・ルールの確立とアーカイブ・システムの設立」が求められると指摘する（小林 2014 : 123）。そして「アーカイブ・ルールの確立」のためには、「まず、調査対象者の許諾を得ることのほかに、調査データへのマスキングなどの処理、非公開期間の設定、公開対象の限定、公開手続きの明確化、これらの手続きの方針を確かにすることが求められる」（小林 2014 : 123）と具体的な事項を挙げてルール確立の必要性について述べている。社会調査データの公開・二次利用に関しては、「日本社会学会倫理綱領」および「日本社会学会倫理綱領にもとづく研究指針」の遵守を前提とした利用規定の作成をルールに盛り込むことも必要となるであろう。さらに、アーカイブ化に伴い生じるであろう権利処理については、法的基盤が未整備な状態であり、吉見俊哉が指摘するように「今日、まず何よりもデジタルアーカイブ化と再利用が円滑に進むように、権利処理の法的な仕組みが整備されること」（吉見 2014 : 567）が喫緊の課題であると考えられる。「アーカイブ・システムの設立」のほうは、非公開データの期限を担保する永続性の確保されたアーカイブ機関の設立は容易ではない現状にあることから、横断型ネットワークで結ぶ非施設型のアーカイブ構築の可能性について言及されている⁷⁾。

これまで述べてきた研究データのアーカイブ化をめぐる現状と課題をふまえ、関西学院大学先端社会研究所として、今後どのようなデータ・アーカイブを作成し提供していくことが学界や社会に貢献することになるのか、長期の総合的な展望が必要であると考えられる。まずはその取り組みの一環として、当研究所の研究員間における社会調査データのアーカイブ化の意義への共通認識を高め、公開・共有に向けた議論を深めていきたい。

謝辞：「平成 29 年度アーカイブズ・カレッジ（史料管理学研修会通算第 63 回）短期コース」では、アーカイブズの理論と実際について体系的に学ぶことができ、自身の問題意識を明確にする大変有意義な機会をいただきました。ここに記して感謝申し上げます。

7) 小林多寿子がフォーラム「質的調査データの公共性とアーカイブ化」における討論での平野健一郎氏の提案に示唆を受けて記述したものである。小林多寿子「質的調査データの公共性とアーカイブ化の問題」『フォーラム現代社会学』13 巻、2014 年、p.123。(DOI: https://doi.org/10.20791/ksr.13.0_114)

参考文献

- 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(0): 103-112.
- 武田尚子, 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』No.665: 71-80.
- 前田幸男, 2017, 「社会調査個票デジタルデータの収集、保存、二次分析について」14th International Conference on Digital Preservation. Kyoto, Japan, September 25-29, 報告資料。ipres2017.jp/wp-content/uploads/jtutorial_maeda.pdf (2017年12月11日アクセス)
- 石田佐恵子, 山田富秋, 2014, 「特集『映像アーカイブズを利用した質的調査の探求』によせて」『社会学評論』65(4): 454-464. (DOI: <https://doi.org/10.4057/jsr.65.454>)
- 桜井厚編, 2014, 「〈特集II〉質的調査のアーカイブ化の問題と可能性」『フォーラム現代社会学』13: 111-113. (DOI: https://doi.org/10.20791/ksr.13.0_111)
- 森本一彦, 2014, 「大学研究所蔵の社会調査報告書のデータベース化の諸問題」『フォーラム現代社会学』13: 125-132. (DOI: https://doi.org/10.20791/ksr.13.0_125)
- 小林多寿子, 2014, 「質的調査データの公共性とアーカイブ化の問題」『フォーラム現代社会学』13: 114-124. (DOI: https://doi.org/10.20791/ksr.13.0_114)
- 吉見俊哉, 2014, 「知識循環型社会とアーカイブ——知のデジタルターンとは何か」『社会学評論』65(4): 557-573. (DOI: <https://doi.org/10.4057/jsr.65.557>)