

【論文 (Peer Reviewed)】

欧州における国家的な質的データのアーカイブを巡る議論 の検討

日本での取り組みに向けて

日本学術振興会特別研究員 PD/中央大学

伊東 香純

要旨：

日本では、多機関の連携による国家のインフラを使った質的データのアーカイブ機関は、運営されていない。本稿は、欧州の国家的なアーカイブ機構が直面してきた問題とその対応策を、質的データゆえの倫理的問題、データの利用の方法論的問題、アーカイブの技術的問題に分けて検討することを目的とする。検討の結果、欧州では、提供者の同意を取り、匿名化し、アクセスを制限することで倫理的問題に配慮しつつ、研究成果に貢献してきたことがわかった。他方で、語り手が誰かが大切な場合や、記録が多くの人に公開されることに価値がある場合もある。そのようなアーカイブは、各大学や市民社会組織によって日本を含めて各地でおこなわれてきた。それらの機関の十全な活動のためには、大規模なアーカイブの存在が重要である。欧州では、国のインフラや学術機関の資金の利用によって、大規模なアーカイブの技術的問題を解消している。日本でこのようなアーカイブを構築するための1つの解決方策は、日本学術振興会の資金を使って学会が運営を担うというものだ。

キーワード：

アーカイブ、質的データ、欧州、研究倫理、二次的分析

1. はじめに

質的データのアーカイブは、特に大規模なものは、西洋の多くの地域において、量的データのアーカイブに遅れて1990年代終盤から始まった。「質的データのアーカイブ化と再利用を推進するために作られた世界初の組織」(Heaton [2008:37])は、英国の「クオリデータ (Qualidata)」であるとされる。クオリデータは、最初にエセックス大学に設立され、その後、全国的な質的データアーカイブに発展し、さらに、量的データも合わせた全国アーカイブに接続されていった(青山 [2019:97])。欧州の中では、英国、アイルランド、フィンランドが、質的および質的縦断的データのアーカイブのインフラという観点でもっとも発展した地域とされている。これらの地域は、国家が社会科学のデータのアーカイブのための資金を提供しており、アーカイブに関わるさまざまな機関のネットワークが形成されているという特徴がある (Neale and Bishop [2011:7])。

日本でも、質的データをアーカイブする取り組みは、数多くおこなわれてきた★01。しかし、英国、アイルランド、フィンランドのような多機関の連携による国家のインフラを使ったアーカイブ機関が運営されるには至っていない。そこで本稿は、欧州で質的データのアーカイブの機構を構築していくにあたってどのような問題に直面し、それらにどのように対応してきたのかを明らかにすることを目的とする。本稿での検討を基に、日本の既存のアーカイブ機関とは異なるそれらの機関の意義を明確にし、日本でも同様のシステムを構築する際の検討の材料としたい。

英国の質的アーカイブ発展の功労者のビショップ (Libby Bishop) は、質的データのアーカイブに対する反対の理由は、技術的な理由、方法論的な理由、倫理的な理由の3つに大別できるという (Bishop [2009:256])。本稿も、この3つの論点を踏襲して論を進めていく。ビショップによると、英国では、データ共有の技術的なことは、完全ではないけれども着実に進歩している。研究の方法論的な問題は、最初のうちは文脈横断的な議論が続いていたけれど、その後「二次的分析」から「再文脈化 (recontextualisation)」(Moore [2007]) という枠組みへの転換が受け入れられるようになっていった。このように比較的解決されつつある前2者に対して、大きな問題として残っているのが倫理的問題であるというのがビショップの見方である (Bishop [2009:256])。

2. 倫理的問題

質的データのアーカイブに際して、もっとも大きな問題として残っているとされる倫理的問題について、具体的にはどのような懸念が表明され、それに対してどのような解決策が講じられているのかを見ていく。パリー (Odette Parry) とモースナー (Natasha Mauthner) は、量的データのアーカイブは問題ないが質的データのアーカイブにはその研究者から複雑な反応があるという状況について、英国社会学会の雑誌『Sociology』で説明している。このような懸念の背景には、英国における質的研究の資金提供者である「経済社会研究会議 (Economic and Social Research Council: ESRC)」が、アーカイブを強く求めている状況があった。ESRCは、助成金を受けた研究には、量的研究でも質的研究でもデータをアーカイブ機関に預けるように求め、それが履行できない場合には経済的な罰を受けることもあり得るとした。その時点で、アーカイブを資金提供の条件としている機関は、ESRCのみであったが、アーカイブすべきという強い風潮があり、他の慈善の団体もアーカイブを推奨している状況であった。このような状況を受けてパリーとモースナーは、クオリデータの実践に焦点を当てて、アーカイブされた質的データは、量的データと同様に、新たな本質的な発見や理論を生み出すために利用できるという前提を批判的に検討した (Parry and Mauthner [2004:140-141])。

第一の懸念は、データの所有権である。パリーとモースナーは、インタビューを録音する場合には、録音データと文字データでは著作権の所在が異なり、前者ではインタビューをコーディネートした機関、後者では話者にあると考える。もちろんこの著作権は、移動可能ではある。例え

ば研究者は、口頭での同意のみで、協力者に説明した目的に沿ってデータを利用する場合がある。しかし、アーカイブを考えた場合、どのような「利用」が将来的に想定されるか、「著作権デザイン特許法(1988)」には明確に規定されていないという問題がある。口頭での同意は、法的には弱い位置づけにあり、データのアーカイブは、書面での同意を取るよう研究者に強く迫るものとなる。さらに、パリーとモースナーは、著者の人格的利益を保護する著作者人格権は、インタビュアー、インタビュイー、調査の資金提供者がデータのコントロール権を放棄した後であっても継続するという。これは、量的データにも質的データにも当てはまるが、質的データに関しては一層重要性が増す事項である(Parry and Mauthner [2004:141-143])。

第二の懸念は、内密性(confidentiality)である。ここには、調査協力者の匿名性と研究者の匿名性が含まれる。調査協力者について、その個人情報内密にされ、匿名性が保持されることが英国を含めて多くのアーカイブで明記されている。2000年3月に施行された英国の「データ保護法(1998)」は、個人のデータの取得、保存、利用、公開について定めた法律である。この法律では、研究を目的とする場合、一定の規定を免除している。しかし、この免除は、そのデータが特定の個人に影響する方法で使われたり、「重大な危害」を生じさせる恐れがあったりする場合には、適用されない。また、人種、政治的主張、健康、性生活、犯罪記録等の特定のデータは、他のものと比較して取扱いに注意が必要とされている。法律では、一次データについてもアーカイブされたデータについても、協力者がその人に危害を与えうるデータ処理の仕方に異議を唱えられるよう保証することを推奨しているものの、実際には研究者コミュニティに公開してしまったデータについて協力者の利益を保護することは難しい。加えて、社会学の質的データの中には、協力者のコミュニティが簡単に特定できてしまうような詳細な民族誌調査や、特定できないように加工してアーカイブすることにより価値が大きく減ってしまうデータが存在する(Parry and Mauthner [2004:143-144])。

次に研究者の匿名性について、アーカイブは、研究者が可視化されにくい実証主義の量的調査から始まっているため、言及されることが少ない。他方で、質的研究では、研究者と協力者が共同でデータを生産し、その過程で研究者が自身の個人的な情報をさらすことが少なからずある。自身の経験を語ることは、研究者自身の経験が研究テーマと接続している場合、協力者とのラポール形成に役立ったり、研究者と協力者の非対称な関係を少しでも平等に近づけたりといった効果をもつ。しかし、研究者は、研究結果を発表する過程で自分の名前等を公表するため、内密性の保持が難しい。質的データのアーカイブの目的の1つは、研究プロセスの可視化であるとされている。この目的の重要性は認められるものの、実践的には研究の生産性を減じる可能性があるという(Parry and Mauthner [2004:145-146])。

第三に、協力者の同意の問題がある。質的研究については、調査協力のインフォームドコンセントが本当に「インフォームド」なものであるかについて長い議論がある。この議論の主たる論点は、質的研究には協力者とのやりとりから得たアイデアを利用しながら調査を進めていく再帰的な性質があるため、事前に協力者に同意を得る段階で明確な研究の流れを説明することが現実的ではないという点である。データをアーカイブする場合、そのデータは一次データを収集す

る研究者の考えなかった利用のされ方をすることになる。パリーとモースナーは、そのための同意をとるのは、第二次利用者が協力者と連絡を取れない場合は特に、一次データを収集する研究者の能力を超えると考える。さらに、研究結果が、協力者の歓迎するようなものでないものになることがありうる。アーカイブしたデータの場合、研究結果が一次データをとる研究者も協力者もコントロールできないものとなる。このようなアーカイブによる影響の予測は、協力者にとっては非常に困難である (Parry and Mauthner [2004:146-147])。

パリーとモースナーは、英国のデータ保護法と ESRC の、特定の協力者は他の人と比較して傷つきやすく、特定のデータは他のものよりも取扱いに注意を要するという立場を 2 つの理由から批判している。1 つは、データの所有権を基本的なこととして扱っていない点である。つまり、傷つきやすい立場の人に限らずすべての研究協力者は、データの利用あるいは再利用の仕方が不適切あるいは軽蔑的な (derogatory) ものだと感じた場合には、異議申し立てする権利を持っていることを十分に認識できていないという点である。もう 1 つは、誰の立場からどのような基準で取扱いに注意を要するデータと判断しているのかが問題となる点である。特定の協力者を特権的に扱うことによって、特定の研究者のデータアーカイブに対する異議申し立てが、他の研究者のものと比較して正当性を認められやすくなってしまふ (Parry and Mauthner [2004:147-148])。

質的研究の中でもオーラルヒストリーの方法を採る研究者は、歴史的な記録の保存を目的として、アーカイブを当然のこととして受け入れてきた傾向がある。他方、他の社会科学の質的研究では、データを使用する際の認識論的な問題がより重要となる。オーラルヒストリー研究にとってデータの収集は目的そのものであるのに対し、他の社会科学の質的研究ではデータの収集は目的のための手段に過ぎない。このような違いを述べた上で、パリーとモースナーは、質的研究者がオーラルヒストリーなどの他の分野の経験を学ぶべきであると主張している (Parry and Mauthner [2004:148-149])。

パリーとモースナーの懸念に対して、同じく『Sociology』上でビショップから、すべての懸念はクオリデータによって既に対処されているとの応答があった。ビショップは、質的研究における研究者と協力者の特殊な関係性ゆえに、1 つのガイドラインをすべての研究に適用することはできず、学会等のガイドラインや倫理規定を参照しつつ、データを預けようとする研究者と会話することで、問題点を解決すべきだと考えている。

個人情報の保護のための匿名化についてビショップは、匿名化によってデータの意味がほとんどなくなってしまうことや、匿名化が十分にできないデータがあることを認める。その場合には、データを受入れていないという。しかし、匿名化はデータの保護のための方法の 1 つに過ぎず、クオリデータでは偽名を使ったり文章を削除したりといった加工の仕方について、研究者に支援をおこなっている (Bishop [2005])。

同様に、アイルランドの「アイルランド質的データアーカイブ (Irish Qualitative Data Archive: IQDA)」では、質的データのアーカイブに関する倫理的懸念を踏まえて、アーカイブのベストプラクティスについてのハンドブックを子ども発育機構と共に作成した。これは、預け入れの同

意書、望ましいアーカイブのフォーマットといった3つの付録まで含めて17ページの冊子である。ハンドブックでは、アーカイブの手続きを6つの側面に分けて説明している。具体的には、データ管理の計画、道具とソフトウェア、データの準備、データの処理、データの預け入れ、データの配布である。それぞれの側面についての注意点が、データを集める研究者とアーカイブを管理する管理者に対して書かれている。

まず、データ管理の計画では、データの取り扱い注意の度合いを、個人が特定される可能性の高さと特定された場合の危険性の高さによって、評価するようになっている。その結果によって、取り扱いに注意が必要な部分を削除するか否かの判断が変わってくる。基本的に、アーカイブの同意と匿名化は必須で、データのアクセスには制限が必要という方針である。データの準備と処理では、匿名化と削除の方針について説明されている。例えば、ジョンという名前はピーターと匿名化されるべきであるが、ロンドンというジョンの出身地は十分に広く特定の危険が低いことからロンドンのままでよいといったことである。自動で匿名化するツールや、匿名化の際に起こりやすいミスなども紹介されている。また、特に重要なこととして、匿名化あるいは削除をした場合には、どこからどこまでが編集した部分なのかを明記することに言及されている。データの配布では、データのアクセスを提供する前に、利用者の認証をおこなう手続きが必要だと述べられている (Gray et al. [2011])。

クオリデータの設立に貢献したコーチ (Louise Corti) らは、すべてのデータが個人の特定を可能にする箇所の削除を必要とするわけではないと述べる。その上で、匿名化以外のデータの保護の方法として提示されているのが、アクセスの制限である (Corti et al. [2000])。クオリデータでは、「匿名性の高さによって「公開データ」(open data)、「保護データ」(safeguarded data)、「管理データ」(controlled data)の三つの段階に分けられ」ている。2019年時点で「約7600件あるデータセットのおよそ9割が『保護データ』で」ある。さらに、「『保護データ』と『管理データ』は非商業目的なら誰でも利用申請できるわけではなく、『学生』まで、『専門的な学生』まで、『研究者・政策決定者』のみ、あるいは研究者に直接コンタクトをして許可を得た者など、場合により利用者がさらに制限されてい」る (青山 [2019:100])。IQDA に所蔵されているデータも、中には誰でも閲覧できるものもあるものの、大半の閲覧は、研究者、教員、第3段階の学術組織に登録している学生に限られている。さらに、国家的なデータ公開のネットワークを介してではなく、IQDA に直接申請書を郵送する必要のあるものもある。(伊東 [2022])。フィンランドの「フィンランド社会科学データアーカイブ (Finnish Social Science Data Archive archives: FSD)」では、データ公開の条件に4つの段階を設けている。具体的には、すべての人がアクセスできるもの、研究、教育、勉強のためであればアクセスできるもの、修士課程以上の研究者ならアクセスできるもの、データを作成したあるいは預けた人の許可があった場合にのみアクセスできるものである (Khan et al. [2021:4])。

さらに、ビショップは、一次データの取得者が預けたデータをコントロールできなくなるというのは誤解だと説明している。クオリデータでは、データをダウンロードできるのは、システムに登録した利用者のみであり、一次データをとった研究者が利用目的を制限したり、自身が許可

した場合にのみデータを利用できるようにしたりすることもできる。また、アーカイブが、研究プロセスを監視するような役割を果たし、非生産的な結果を生むのではないかという懸念に関しては、研究プロセスの透明性が増すことと、それが監視的な役割を果たすことは、分けて考えるべきだとビショップは主張する。クオリデータは、「監視のような複製物」としてのデータ利用を奨励していない。質的方法の発展は、研究者が自身の研究は不十分で公開するに足りないと思っているときにこそ起きるものである。プロセスの透明性が、研究者の深い専門性と思慮深い判断を可視化する。資金提供者にとっては、自身の提供した資金から最大限の成果を得ることにつながる (Bishop [2005])。

パリーとモースナーが挙げた問題点は、長期的に対処していくべきものであるとビショップは最後に述べている。2005年時点でクオリデータは、年間100件ほどのデータを処理しており、その中でアーカイブに選ばれるものは多くないという。分野に沿ったアーカイブのガイドラインを作成していくべきとの意見に同意し、クオリデータがその過程に重要な役割を果たす意思があることを表明している (Bishop [2005])。このようにクオリデータでは、研究協力者を保護するための方法として、同意を得て、匿名化し、データへのアクセスを制限するという基本的な方針が定められていた。加えて、研究者や研究機関の保護や研究活動の生産性という観点から表明された懸念は、アーカイブはそれらに否定的に影響する仕方での利用を奨励していないという反論がなされた。

他方で、質的データのアーカイブは、研究参加者を保護する観点から躊躇される場合が多いが、研究者と研究参加者では倫理的問題に対する認識が異なっていることを示す研究がある。フィンランドのFSDは、2003年に質的データのアーカイブを開始した。その際、多くの研究者は、この取り組みを倫理的に問題があると考えたという。特に、インタビューデータのアーカイブに反対した研究者は、インタビューという状況の内密性を強調した。研究者は、「インタビューは非常に親密で、取り扱いに慎重を要し、予測性がなく、感情的なもの」だと考えていた。そのため、そのフィールドにいなかった人が再利用するための、データのアーカイブは現実的ではないと主張した (Kuula [2011:12])。

インタビューのアーカイブに対する研究者の態度は、研究領域によっても異なるという。フィンランドでは、特に1980年代以降、量的研究よりも質的研究に好意的な傾向がある。社会科学と比較して人文学では、発表された研究の解釈や結果をチェックしたい場合のための証拠としてデータをみなす。加えて、文化や歴史を理解するために保存されるべき、貴重な共有資源とみなされている。これに対し社会科学では、データをより私的財に近いものとみなす傾向がある。法的な側面についても、同様の差異があるという。人文学では、研究参加者の著作権の重要性を強調するのに対し、社会科学では、この傾向は変わりつつあるものの、プライバシーの問題を強調する (Kuula [2011:12-13])。

FSDでは、インタビューイにとってデータのアーカイブは受け入れがたいという想定を検証するための研究を実施した。具体的には、数名の研究者に、その人の研究の参加者に再度連絡を取らせてほしいと依頼し、その研究者とともにデータのアーカイブに同意できるか手紙で尋ねた。

さらに、その中の何人かの参加者には、電話でデータの将来的な利用について質問した。対象となったのは、4つの質的研究で、そのうち3つはインタビュー、1つは大学生のライフストーリーの記述を利用したものであった。すべてのデータは、唯一無二で個人的な語りを含み、中には傷つきやすい話題を含むものもあった。4つの研究の参加者のうち、再度連絡がとれたのは169名であり、そのうち98%にあたる165名が自身のデータのアーカイブに同意したという。さらに、電話での質問によって参加者は、自分のインタビューは研究に足るものであると考えており、科学の発展を主たる理由としてアーカイブに同意していることが明らかになった。参加者の中には、最初の研究の結果に納得しておらず、異なる専門の研究者に再度分析されることを歓迎した人もいた。この研究によって、研究者と研究参加者では、研究の関係性の理解が大きく異なっていることが明らかになった。この結果は、参加者のアーカイブによるリスクを予測する知識の不足によると解釈することもできるが、参加者はインタビューの関係性を私的で内密なものとはみなしていないとも解釈できる (Kuula [2011:15])。

3. 方法論的な問題

本節では、質的データのアーカイブを利用して研究をおこなう際の方法論的な問題について、まず実態としてデータを二次的に分析する研究がどの程度、またどのようにおこなわれているのかを実例を挙げながら説明する。その上で、そのような研究方法を選択する際の懸念と、問題を解決するための方法をみていく。

ヒートン (Janet Heaton) は、質的データの二次的分析について、文献レビューをおこなった。二次的分析とは、先行研究のデータを再利用した分析によって、新たな問いに答えたり、先行研究の知見の妥当性を検証したりすることである。二次的分析に利用するデータの取得の仕方は、アーカイブ機関などからデータを得る、インフォーマルなネットワークを通してデータを共有してもらい、自分で集めたデータを再度分析する、の3つに分けられる。ヒートンは、二次的分析が最もよく用いられていると考えられる「保健と社会的ケア」の分野の論文を調査した結果、65件の二次的分析を用いていると考えられる論文を発見した。そのうち41件は、二次的分析を用いていることを明記していた。二次的分析に用いられる一次データの78%は、北米と英国の研究によるものであり、二次的な研究もフランスなど他の欧州地域で実施されていた (Heaton [2004:36-37])。

65件の86%にあたる56件の二次的な研究では、1人以上の筆者が一次データの収集に関与していた。残りの14%にあたる9件の研究のうち、2件だけが公的にアーカイブされたデータを使った研究であった。そのうちの1件がブルア (Michael Bloor) による研究ノートである (Heaton [2008:39])。この研究は、医療社会学で分析されてきた医師-患者関係に関する実証的研究を、科学知識の社会学の理論的な枠組みから検討するというものである。具体的には、科学実験に対する一般社会からの影響、論争的な知識、科学知識の普及に注目して、鉱山労働者の賠償請求に

関するオーラルヒストリーを分析する。データは、英国のウェールズ・スウォンジー大学の「南ウェールズ鉛山労働者図書館」に所蔵されている、いくつかの個別の研究で構成された「炭鉛の歴史プロジェクト」のものを用いた。インタビューの選定の方針は、研究によってさまざまであった。ブルアは、561件のインタビュー音源データのうち、1973年と74年に「南ウェールズの炭鉛の歴史プロジェクト」の一部として収集された176件のデータを利用した。分析の結果、炭鉛労働者と医学的科学的知識との間には、ザイマン (John Ziman) が「道具的」と呼んだ関係があることがわかった (Bloor [2000])。このようにブルアは、一次的な分析とは異なる理論枠組みを用いた二次的な分析により新たな知見を得ることに成功している。

2004年の書籍 (Heaton [2004]) 出版以降、発表された公的な質的データのアーカイブを利用した研究として、ヒートンは2件を挙げており、そのうち1件がSavage (2005)である (Heaton [2008:39])。この研究では、英国の労働者階級に関する研究を対象に、それらの研究を3つの世代に分けた上で、特に1950年代から1970年代の第一世代の研究に注目する。それらの研究者によるインタビューを再度分析することにより、第一世代の理論が1970年代になぜ勢いを失っていったのかを事例検討として明らかにする。第一世代の理論について既に指摘されてきたことの妥当性を認めて繰り返すことに加えて、インタビューデータを今日的な視点から分析することにより、当時の研究とは異なる指摘をおこなう (Savage [2005:930])。分析に使用したデータは、エセックス大学のクオリデータに所蔵されている227件のインタビューである。第一次の分析では、金銭的な点など量的データに変換しやすい側面が注目され、質的データが研究中に十分に引用されていないのではないかと考え、第二次の分析ではインタビューデータを引用しながら分析した (Savage [2005:932-933])。その結果、現在の社会学の知見から見ると、何が「普通」なのかはインタビューによって異なっており、インタビューの認識に基づいて社会階層について考察するという方法には問題がある。このような理由で、第二世代の階級研究者が個人の認識に注目するのをやめていったという流れが納得できると結論付けている (Savage [2005:942])。この研究では、先行研究の妥当性を検証するとともに、当時は予想できなかった研究の歴史を踏まえた上で、その背景を検討している。

クオリデータによる1999年の調査では、過半数の研究者がデータアーカイブにアクセスしたいと考えていた。しかし、その後の英国の「経済社会研究会議 (ESRC) のデータ取り扱い方針とアーカイブ」によると、研究業界にはデータのアーカイブと二次利用について、3点の懸念がある。第1の問題は、第一次の目的のために収集されたデータが、第二次の目的のための分析に利用するデータとしての的確であるのかという点である。第2の問題は、適切な同意のとり方や秘密の守り方といった倫理的、法的な問題である。第3の問題は、質的データのアーカイブや再利用に関する政策やガイドラインの不足である (Heaton [2008:40-41])。

ヒートンが指摘した懸念のうち第1の問題に関して、フィールドイング (Nigel Fielding) は、質的データは、研究者と協力者のやりとりを通して生み出されるものであり、歴史的政治的認識論的制約を受けているため、方法論的な探求以外の目的での二次的な分析は有効ではないという批判を取り上げている。このような批判に対してフィールドイングは、質的研究では再帰性の影

響の分析と考慮が常に重要なので、文脈の影響の評価に関して、一次データの分析と二次データの分析の間に論理的な不一致はないと反論する。しかし、認識論的な問題よりも重要なのは、実践的な問題であるという。つまり、データの豊かさが質的研究の強みであり、分析において重要なデータが書き落とされてしまうことは、一次データでも二次データでも同様に問題といえる。フィールドワークの途中で調査の焦点が変化することは往々にしてあり、そのような変化は確実にデータに影響を及ぼす(Fielding [2004:99])。

さらに、一次データの収集者は特定の問題関心を持ってフィールドワークを始める。その問題関心が、データを歪める効果を持っていることは従来から言われており、研究者がこの効果に自覚的でない場合には、その影響はいつそう顕著になる。このような影響を抑えるために、フィールドワーカーが他の研究チームのメンバーにデータの分析を依頼することがある。アーカイブされたデータの二次分析は、これと同様の役割を果たすことができるというメリットがある。フィールドワーカーは、二次分析の実施を頭に置きながら調査する必要がある(Fielding [2004:100-101])。

かつてはデータの収集に関わる点が、質的研究の妥当性の判断において重要な位置を占めていた。例えば、どの程度の期間フィールドで過ごしたかといった基準である。しかし、近年より重要性を増してきているのは、「転用可能性」「信頼性(credibility)」「信頼性(dependability)」と呼ばれるものである。これらは、批判的な査読、分析の妥当性についてのコミュニティにおける基準、経験からの学習によって判断される。この場合、データの背景はそれらの判断の材料として使われる(Fielding [2004:102])。加えて、データをアーカイブ機関に預けることにより、貴重なデータをよい状態のまま安全に保管しておける可能性が高くなる。さらに、著作権や内密性といったかつてと比較して近年重要性が増している質的研究の研究倫理についても一定の基準を満たすことができる。このように専門の機関によるアーカイブは、一次データを収集した研究者が二次的分析を実施する場合にも利益になる(Fielding [2004:103])。

フィールドイングは、アーカイブされたデータの使い方は、質的研究の間でも分野によって異なっており、そのため「アーカイブデータの再利用」と「データの二次的分析」は分けて考えた方がよいと述べる。アーカイブデータの再利用は、時間的に再度利用することを指し、特定の事項に関して当該データ以外のデータが存在しない可能性があることを示唆する。この場合、アーカイブされたデータの利用は自明に意義があるといえる。他方で、データの二次的分析の場合、同様のデータを同じ時代に収集できることが想定されている。このため多くの研究者は、データの存在論的、認識論的な問題に関する議論に巻き込まれるのを回避するために、一次データを自ら収集し、その分析に力を注ぐ。二次的分析は、歴史研究よりも社会科学の質的研究に当てはまるものである(Fielding [2004:104])。

続いて、誰がどのような目的でアーカイブされた質的データを利用しているのかをみていく。「UK データサービス」の、2002年から2016年までの7155回の質的データのダウンロードについて、41.68%にあたる2982回が大学院生、26.86%にあたる1922回が高等教育機関の職員、24.96%にあたる1786回が学部生によるダウンロードであった。それらの7155回のダウンロー

ドの目的について、63.72%にあたる 4559 回が学習 (learning)、15.00%にあたる 1073 回が研究、13.38%にあたる 957 回が教育 (teaching) を目的としていた。学習と研究の区分について、携わっているプロジェクトの内容が妥当性のある研究であれば研究、宿題等の課題を遂行するためであれば学習に分類する。博士号のプロジェクトは研究、学部と修士号のプロジェクトは学習に自動的に分類している。学習目的のうち、57.84%が大学院の学習、38.80%が学部の学習、3.36%がその他の学習であり、教育目的のうち、31.56%が大学院の教育、22.78%が学部の教育、45.66%がその他の教育であった。フィンランドの FSD は、2014 年に「Aila」というオンラインのデータサービスを開始し、その頃からデータの再利用が増加した (Bishop and Kuula-Lummi [2007:4-5])。FSD では、2015 年から 2018 年までの 4 年間で量的データと質的データを併せて、10346 件のダウンロードがあった。そのうち 14.6%が質的データ、85.4%が量的データのダウンロードであった。ダウンロードの目的は、多い方から勉強 (study work) / エッセイが 39.0%、研究が 22.1%、学部及び修士論文が 21.7%、教育が 9.2%、博士論文が 5.0%、その他が 2.9%と続く。量と質で比較した場合、その他の目的を除いて、両者の割合が最も近かったのが学部及び修士論文を目的としたダウンロードで、量的データが 72.9%、質的データが 27.1%であった。他方、両者の差が最も大きかったのが、研究を目的としたダウンロードで、量的データが 94.5%、質的データが 5.5%であった (Late and Kekäläinen [2020])。アーカイブされた質的データの利用の方法論的な問題の議論は、主に研究を目的とした場合に生じる問題を論じている。しかしながら、もちろん研究と学習や教育は密接に結びついたものではあるものの、研究以上に学習や教育を目的として利用されていることがわかる。

4. 技術的な問題

クオリデータは、1994 年に ESRC から 70 万ポンドの助成金を受けて開始した (Corti [2000:para. 43])。青山薫によると、ESRC は「日本でいえば学術会議の社会科学部門」にあたる (青山 [2019:97])。1996 年には、クオリデータは、慈善団体から 5 万ポンドの寄付を受けたほか、ESRC から 40 万ポンドが補充された。1999 年には医学研究会議から 35000 ポンドの助成があった。ESRC からの 5 年間の助成金は 2000 年で終わった (Corti [2000:para. 43])。そして、「1999 年から 2000 年の ESRC など出資機関の予算削減の標的にされ、職員と予算を大幅に削られ存続の危機に立たされ」た (青山 [2019:97])。

1994 年から 2000 年までの 7 年間のクオリデータの職員数について、もっとも多いときには 10 名がいた。しかし、資金が減った 2000 年には、5 名のほぼパートタイムの職員になってしまった。職員には 9 つの役職が設けられている。期間によってポストが埋まっていた時期もあればそうでない時期もあり、そのポストの仕事にかけられていた時間も異なる。9 つのポストは、理事長、副理事長と管理者、管理者、行政担当者、(データ) 取得担当者、研究者支援担当者、処理担当者、事務 / 書記職員、スキャン担当職員である。もっとも職員数が少なかったのが 1994

年で、理事長が 0.4FTE★02、行政担当者が 1FTE、取得担当者が 0.8FTE、処理担当者が 0.5FTE の合計 2.7FTE であった。もっとも多かったのが 1998 年で、理事長が 0.4FTE、副理事長と管理者が 1FTE、取得担当者が 1FTE、研究者支援担当者が 0.5FTE、処理担当者が 2FTE、事務／書記職員が 0.5FTE の合計 5.4FTE であった (Corti and Backhouse [2000])。

5 年間の大型の助成の後、2000 年 10 月から 2001 年 9 月まで、ESRC から 1 年分の助成として 10 万ポンドが提供された。同じく 2000 年、エセックス大学から 12000 ポンドの支援があった。経済的な苦境に立たされたクオリデータは、同じくエセックス大学の UK データアーカイブと合併することで、技術的な職員と設備のアクセスを得るという長期的な戦略を立て始めた。この戦略は、質的データにまでアーカイブの幅を広げ、社会科学者のデータのワンストップ窓口になるという UK データアーカイブの当初の目論見の実現にもつながると考えた (Corti[2000:para. 43-46])。実際にクオリデータは、2001 年に「UK データアーカイブに吸収合併されることで、存続の道を得」た。「量的データ、視聴覚データもふくむ質的データ、混合的な方法でなされた調査のデータを扱っていた専門家たちがお互いのデータ取り扱い方法を学びあい、アーカイブの統合が成し遂げられた」(青山 [2019:99])。

その後、2003 年 1 月に、全国的な社会科学に関するデータサービスとして、「経済社会データサービス (ESDS)」が始まった。経済社会データサービスは、エセックス大学の UK データアーカイブと UK 縦断的研究センターと、マンチェスター大学のセンサス調査研究センターとマンチェスター情報関連サービスの機関が協働で運営している。もともと UK データアーカイブがサービス全体の管理を担当し、クオリデータを運営していた。資金面では、データのアーカイブおよび利用に関する方針の見直しをおこなっていた ESRC が新たな助成を決めた。そして経済社会データサービスが、ESRC と、高等教育にデータサービス等を無料で提供するための高等教育資金会議 (HEFC) の共同情報システム委員会 (JISC) の資金提供で始まった。これによりクオリデータの永続性が強まった (Economic and Social Data Service [2004:3]; 青山 [2019:99])。

クオリデータでは、発足して数年後に、どこにどの程度の質的データが保存されたり共有されたりしているのかを明らかにするための探索的な研究をおこなった。クオリデータの重要な目的の 1 つが、英国に現在実際にある、及びある可能性のある質的データを特定し、それらを公開することであったからだ。そこで、社会的調査のデータを所蔵していると思われる機関及び個人を対象に、5 つの項目について調査を実施した。対象となった機関等は、図書館などの伝統的なアーカイブ機関、公文書館と博物館、質的研究に従事している学術研究グループ、個人の研究者、データアーカイブとデジタル図書館の 5 つである。調査項目は、収集の根拠、質的データの例、記録媒体と保存の理由、アクセスのしやすさと共有の意思、発見しやすさと宣伝の 5 項目である。本稿では、このうち収集の根拠と、アクセスのしやすさと共有の意思に注目する。収集の根拠は、あるデータを収集するかないかを判断する根拠の有無についてである。根拠がある場合が多いのが、個人の研究者、ない場合が多いのが、図書館などの伝統的なアーカイブ機関、データアーカイブとデジタル図書館である。アクセスのしやすさと共有の意思は、高いのが、図書館などの伝統的なアーカイブ機関、公文書館と博物館、データアーカイブとデジタル図書館で、低

いのが、質的研究に従事している学術研究グループ、個人の研究者は高い人から低い人までさまざまであるという結果になった。このような結果に対してクオリデータでは、収集の根拠を持ち、アクセスのしやすさと共有の意思が高い機関を目指すと述べている (Corti [2000])。このように英国でも、クオリデータより前から質的データのアーカイブはさまざまな機関でおこなわれていた。クオリデータは、これまでのどのようなアーカイブ機関、個人とも異なる機関を目指して設立されたことがわかる。

5. まとめ

本節では、質的データのアーカイブをおこなう際に考えるべき問題に対する、欧州の大規模なアーカイブ機関の取り組みを参照しながら、日本で今後、質的データのアーカイブを進めるために必要なことを考察していく。欧州の取り組みにおいて今後考えていくべきもっとも大きな問題とされていたのが、質的データゆえに強調されたり生じたりする倫理的な問題である。この問題については、複数の観点から懸念が示され、それに対する反論、反証が示されてきた。その反論、反証には一定の妥当性が認められた。倫理的問題の議論に決着が着いたわけではなく、議論の継続が重要だとされている。しかしながら、研究協力者の保護のためにとるべき方針については、概ね統一された見解がある。具体的には、データ提供者の同意を取り、匿名化し、データへのアクセスを制限するという方針である。

アーカイブされたデータを研究に利用する際の方法論的な問題に関しては、既にそのような方法を使った研究成果が発表されていることがわかった。さらに、それらの中には、一次データの収集者とは異なる目的のもとに二次的な分析をおこない、新たな知見の提示に成功している研究もあった。加えて、高等教育機関における教育を目的とした利用も広く行われており、研究者養成という観点からも質的データのアーカイブは、社会科学の研究を発展させる上で意義のあるものといえる。このように、欧州の大規模な取り組みは、倫理的問題に配慮しつつ、研究成果の生産に貢献してきたことが明らかになった。

他方で、匿名化やアクセス制限という方針が適用しにくい研究もある。例えば、UK データアーカイブは、家系研究や家族史研究には向かないとの注意書きがあるという (青山[2019:100])。本稿ではほとんど取り上げられなかったが、ある目的をもってある範囲の記録を集め公開する試みも世界中にある。その場合には、誰が語っているのかが大切な場合があり、またその記録は、多くの人に公開され読んでもらうことに価値があるとされる。そのような質的データを、倫理的な問題に可能な限り配慮した上で、アーカイブしていく取り組みは、各大学の研究所や図書館、市民社会組織などによって担われてきた。これらは、クオリデータや IQDA などと比べるとかなり規模の小さなものである。

そのような小規模な取り組みの中でも比較的規模の大きなものが、「ディベックス (Database of Individual Patient Experiences: DIPEX)」プロジェクトである。このプロジェクトは、臨床的

な根拠の系統的評価のためのコクラン★03の構成部分の1つとして、保健、疾患、ケアに関する人々の経験の一般公開のデータベースを作成したいというアイデアから1999年に始まった。当初は、オックスフォード大学の前立腺癌の人々の経験に焦点を当てた小さな研究グループによる1つのプロジェクトであった。その後、ネットワークが広がっていき、2021年現在では14か国がメンバーカントリーとなり、ウェブサイトは9つの言語で運営されるようになった。国際的な連携により、150種類の状況についてのページが作成済みであり、さらに43種類の状況についてのページが準備中である。毎年、850万人のアクセスがあり、発足から20年間の外部からの資金は全部で3000万ユーロを超える(Zieband et al. [2021:133])。ディベックスのデータは、偽名を使う場合が少なくないが、多くのデータが特定の健康上の問題を持った語り手の映像とともに公表されている。そのデータは、研究利用に加えて、同様の問題を体験している人や家族などの情報源になることを意図して、誰でもアクセスできるウェブサイト上に公表されている。

ディベックスプロジェクトのデータ公開のウェブプラットフォームのうち、英国のものはヘルストークと名付けられている。このヘルストーク上のデータを有効に利用するための研究もなされている。例えば、患者中心のサービス改善アプローチを作成、試運転、評価するために、患者の経験の語りについての全国的な動画、音声のアーカイブを利用してみようを目的とした研究がある。この研究の背景には、患者の経験に基づく医療サービス改善の重要性がある。この重要性は、広く認められてはいるものの、実施に時間と資金がかかることが大きな障壁であった。そこで、アーカイブを利用することにより、その負担を軽減しようというのが、この研究の意図である。通常であれば、その地域の患者の経験を集めて、サービス改善のためのワークショップの議論のきっかけとして使うところを、ヘルストークに掲載されている映像を利用した。結果として、アーカイブの利用により、全体の期間を1年から半年ほどに短縮することができた(Locock et al. [2014])。

クオリデータなどの大規模な取り組みでは対応できないアーカイブが、多くの個人や組織によっておこなわれており、研究という側面でも成果を挙げている。このような取り組みは、日本にもある。収集の目的に応じた多様な取り組みが、質的データであるから一層、目的に応じた方法で継続される必要がある。そのようなアーカイブに携わる個人や組織が、それぞれの活動に専心し、その活動が十全におこなわれるためには、クオリデータのような機関による大規模なアーカイブが重要である。それにより、大規模な機関に任せられるアーカイブと、そうでないものを分擔することができる。

同意を取り、匿名化し、アクセスを制限するという方針を守って質的データのアーカイブをおこなう場合、人手とインフラが必要になる。ここで、生じるのが技術的な問題である。英国やアイルランドでは、個別の研究者の負担を減らして大切なデータを漏れなくアーカイブするために、同意のとり方や匿名化の仕方の相談にのっており、専門的知識をもった人を配置している。このような体制を維持するには、少なくない経済的負担が必要になる。実際に、質的データのアーカイブは、英国ではエセックス大学、アイルランドではメイヌース大学で始まったが、1つの大学

では担いきれなくなり、全国的なインフラを使って運営するようになっていった。クオリデータは、日本学術振興会の社会科学部門にあたる機関と高等教育資金会議の資金で運営される経済社会データサービスの一部となることで、技術的な問題を解消している。

日本の質的データのアーカイブの状況は、各研究機関や市民社会組織による先駆的な取り組みは進められているものの、短期的な個別のプロジェクトとしてではなく持続的な研究インフラとして国の予算やインフラを使って運営されているクオリデータのような機関は存在しない状況である。このような役割を担う質的データのアーカイブを構築しようとする場合、英国などの場合と同様に1つの大学で担いきれるものではないと考えられる。そこで、1つの解決策となりうる方法は、日本学術振興会の資金を使って、日本社会学会などの学会がアーカイブを担うというものだ。これにより、既になされているアーカイブの取り組みの不必要な負担を減らし、貴重なデータをより多くの研究成果に繋げることができる。

■註

★01 日本における質的データのアーカイブの取り組みの中でも社会科学、社会学に関するものについては、立命館大学生存学研究所のホームページ「アーカイブ 於：社会学・社会科学&学会」(<http://www.arsvi.com/a/arc-soc.htm>)に一覧表がまとめられつつある。

★02 FTEは、フルタイム当量 (full-time equivalent) の略である。通常、1週間に1日8時間ずつ5日間働くフルタイム勤務の従業員の仕事を1として換算する。

★03 コクランは、「根拠に基づく医療 (EBM)」の実践のための根拠作りに貢献すべく、1992年にオックスフォード・コクランセンターで始まった。2019年現在では、130か国以上に45,000人の関係者を擁するグローバルな非営利機関として、治療試験の分析データを提供している。コクランの活動の中心は系統的レビューで、コクラン図書館で閲覧できる成果は、臨床家や研究者、保健政策関係者に利用されている。しかし、コクランの重要な目的は、患者やそのケアをする人々に、意思決定をする際に信頼できる根拠を提供することである。このためコクランUKのブログは、コクランのレビューを医療専門家でない人を含めて、多くの人が読みやすいよう工夫されている (Chapman and Ware [2019])。

■文献

青山 薫 2019 「UK 質的データアーカイブの設立経緯とその後」、『立命館生存学研究』3:97-103

Bishop, Libby 2005 “Protecting Respondents and Enabling Data Sharing: Reply to Parry and Mauthner,” *Sociology*, 39(2):333-336

——— 2009 “Ethical Sharing and Re-use of Qualitative Data,” *Australian Journal of Social Issues*, 44(3):255-272

- Bishop Libby and Arja Kuula-Lummi 2017 “Revising Qualitative Data Reuse: A Decade on,” Sage Open, 1-15
- Bloor, Michael 2000 “The South Wales Miners Federation, Miners’ Lung and the Instrumental Use of Expertise, 1900-1950,” *Social Studies in Science*, 30 (1):125-140
- Chapman, Sarah, and Lynda Ware 2019 “Spreading the Word: Evidently Cochrane,” *Maturitas* 124:89-90
- Corti, Louise 2000 “Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research: The International Picture of an Emerging Culture,” *Forum Qualitative Sozialforschung*, 1(3): Article 2
- Corti, Louise, and Gill Backhouse 2000 “Qualitative Archives Short Descriptions: ESRC Qualitative Data Archival Resource Centre (Qualidata), University of Essex, UK,” *Forum Qualitative Sozialforschung*, 1(3): Article 31
- Corti, Louise, Annette Day, and Gill Backhouse 2000 “Confidentiality and Informed Consent: Issues for Consideration in the Preservation of and Provision of Access to Qualitative Data Archives,” *Forum Qualitative Sozialforschung*, 1(3): Article 7
- Economic and Social Data Service 2004 “Economic and Social Data Service Strategic Plan 2004-2009,” (2022年10月12日取得, <https://ukdataservice.ac.uk/app/uploads/esds-strategicplan20042009.pdf>)
- Fielding, Nigel 2004 “Getting the Most from Archived Qualitative Data: Epistemological, Practical and Professional Obstacles,” *International Journal of Social Research Methodology*, 7(1):97-104
- Gray, Jane, Julius Komolafe, Hazel O’Byrne, Aileen O’Carroll, and Tara Murphy 2011 “Best Practice in Archiving Qualitative Data,” *NIRSA Working Paper Series*, 65. (2022年8月16日取得, https://mural.maynoothuniversity.ie/2691/1/JG_Best_Practice.pdf)
- Heaton, Janet 2004 *Reworking Qualitative Data*, London, Thousand Oaks, and New Delhi: Sage Publications
- 2008 “Secondary Analysis of Qualitative Data: An Overview,” *Historical Social Research*, 33(3):33-45
- 伊東 香純 2022 「質的データのアーカイブに関するアイルランドにおける取り組み」, 『遡航』 3:14-26
- Khan, Jawaria, Visa Rantanen, Petteri T Kolmonen 2021 “Open Access to Qualitative Research Data: A Guide for PhD Students in the Social Sciences at the University of Helsinki,” Population Ageing and Its Social Consequences Working Paper No. 9, (2022年9月4日取得, <https://www2.helsinki.fi/sites/default/files/atoms/files/wp9.pdf>)
- Kuula, Arja 2011 “Methodological and Ethical Dilemmas of Archiving Qualitative Data,” *International Association for Social Science Information Service and Technology Quarterly*,

34(3-4):12-17

- Late, Elina, and Jaana Kekäläinen 2020 “Use and Users of a Social Science Research Data Archive,” *PLoS One*, 15(8):e0233455
- Locock, Louise, Glenn Robert, Annette Boaz, Sonia Vougioukalou, Caroline Shuldham, Jonathan Fielden, Sue Ziebland, Melanie Gager, Ruth Tollyfield, and John Pearcey 2014 “Testing Accelerated Experience-based Co-design: A qualitative Study of Using a National Archive of Patient Experience Narrative Interviews to Promote Rapid Patient-centred Service Improvement,” *Health Services and Delivery Research*, 2(4)
- Moore, Niamh 2007 “(Re)Using Qualitative Data?” *Sociological Research Online*, 12(3)
- Neale, Bren, and Libby Bishop 2011 “Qualitative and Qualitative Longitudinal Resources in Europe: Mapping the Field and Exploring Strategies for Development,” *International Association for Social Science Information Service and Technology*, 34(3):6-11
- Parry, Odette, and Natasha Mauthner 2004 “Whose Data Are They Anyway? Practical, Legal and Ethical Issues in Archiving Qualitative Data,” *Sociology*, 38(1):139-152
- Savage, Mike 2005 “Working-class Identities in the 1960s: Revisiting the Affluent Worker Study’, *Sociology*, 39 (5):929-946
- Ziebland, Sue, Rachel Grob, and Mark Schlesinger 2021 “Polyphonic Perspectives on Health and Care: Reflections from Two Decades of the DIPEX Project,” *Journal of Health Services Research and Policy*, 26(2):133-140

An Examination of the Discussion of the National Qualitative Data Archive in European Countries

Toward Development in Japan

Kasumi Ito

Abstract:

An institution for qualitative data archive that uses national infrastructure and is operated by multi-institutional collaboration has yet to be established in Japan. The purpose of this paper is to describe the problems that national European archival institutions that are already in operation have faced and how they have responded to them. The paper examines three issues: ethical issues due to qualitative data, methodological issues of re-using data, and technical issues of archiving. The results reveal that the European archives settle ethical issues through policies for obtaining research participants' consent, anonymization, and restriction of access, and contribute to researches. However, these policies are not suitable for all studies and data. In such cases, it may be important to note who has provided the data or to make data available not only to researcher. Such data have been archived by various universities and civil society organizations, including these located in Japan. Operation of the large-scale archive is important in order for such small archives to be able to concentrate on their original activities. They large-scale archives have solved technical problems by using a national infrastructure and funding from national academic institutions. To address the technical problem of building a large-scale archive that fulfills such a role in Japan, this paper proposes one solution that academic societies operate the archive by using funds from the Japan Society for the Promotion of Science.

Keyword:

archive, qualitative data, Europe, research ethics, secondary analysis