

❁ データ収集の仕方

International Relations Self-Study Navigator

データ収集の仕方 目次

1. はじめに
2. データ収集以前
3. 二次データから一次データまで
4. 定量的データと定性的データ
5. 文献・データの探し方
6. おわりに
7. 参考となる文献

1. はじめに

ここでは、国際関係学部の1～2回生に向けて、大学での学修におけるデータ収集の意義と方法を述べます。3～4回生にも卒業研究に取り組む際に再確認してほしい内容となっています。

ところで、次のような経験はありませんか？ 「基礎演習」クラスの報告に際してテキストを「熟読」して準備したが、発表後の議論は期待したよりも活気を欠いたものとなった——。教員から「内容を批判的に掘り下げてみよう」とか「発展的に深めてみよう」とかアドバイスをもらったが、具体的に何をしようかわからなかった——。これらに似た戸惑いを感じたことのある人は、普段何気なくおこなっているデータ収集のプロセスに意識を向けてみることで、自分なりの学修・研究スタイルをよりよくするヒントが得られるかもしれません。

2. データ収集以前

あなたの疑問や関心がデータ収集の出発点です。1回生時ならば、基礎演習のテキストに記載されている事項について関心を持つこともあるでしょうし、報告班の仲間と雑談しているうちに疑問が生まれてくることもあるでしょう。それらは、具体的な疑問の場合もあればたとえば「A国の森林面積は近年どのように変化しているのだろうか？」「B国の難民認定者数は増えているのだろうか？ 減っているのだろうか？」など、もっと漠然とした興味や感想かもしれません——たとえば「C国の政府開発援助の使途が気になる」とか、「・・・と書かれて（言われて）いるが本当だろうか？ 納得できない」など——。当然ですが、手元にあるテキストを何度読んでも、記載された情報以外はそこから得ることはできません。自分で考えるにしても、これまでに得た知識だけでは限度があるでしょう。そこで、疑問に答え、関心を満たすために、他の情報源からデータを得て、それらを材料にして考えることが必要となります。

とはいえ、たとえば熱帯林保全の現状に興味をもったからといって、いきなり東南アジアや南米へ飛び立っている人はほとんどいないでしょう。多くの場合は（あるいは現地へ出発する前には）、情報収集の手はじめとして、既存のまとまった資料から調べていきます。つまり、特定のテーマについて論じられた文献—図書や論文、あるいは新聞や雑誌の記事など—から調べていきます。目的とするテーマに関連する文献タイトルをリストアップするにはGoogleの提供する学術情報の検索サービスである「Google Scholar」や国立情報学研究所が提供する検索サービス「CiNii」などが役立ちます（検索時のコツや注意点は後半「5. 文献・データの探し方」で述べています）。また、立命館大学図書館ウェブサイト内の「RUNNERS」を使えばリストアップした文献の所在がオンラインで確認できます（「CiNii」と「RUNNERS」は機能統合しているので、大学図書館

ウェブサイトで両方を同時に使用できます)。その上で、まずは衣笠キャンパス内の図書館にあるものやウェブ上で参照可能なものなど、身近な資料から実際に手にとって目を通してみましょう。さらに、必要に応じて、文献収集の範囲を拡げていきます。他キャンパスや他大学に所蔵されている文献を取り寄せるなど収集範囲を物理的に拡げたり、日本語文献に偏っていたならば外国語（英語）文献も対象に加えたりします。

ひとつひとつの文献に時間をかけて読み込むことが研究をすすめる上で大切なのは当然ですが、学修段階や目的によっては数多くの文献に触れることも必要となります。さまざまな研究者の見解や多くの著者が問題視している事柄を広く知ること、自身の関心や疑問がより明確になるでしょう。たとえば、熱帯林保全への漠然とした興味から関連する文献を調べていくうちに、「社会林業」や「アグロフォレストリー」などの具体的な森林保全活動に興味を持つかもしれませんし、「熱帯林の多様な価値」についてより深く知りたくなるかもしれません。データ収集以前に、あるいはデータ収集をしながら、関心テーマの絞り込みあるいは具体化も同時にすすめます。

3. 二次データから一次データまで

論文や書籍からは整理された状態のデータを得ることができます。個々の論文の著者は、自身の意見を述べると同時に、その根拠となるデータを必ず示しているからです。たとえば、熱帯アジアの途上国における森林消失の深刻さや保全の必要性を主張すると同時に、地域ごとの森林面積の年変化などを根拠として提示しているでしょう。データは、著者自身が現地調査などによって研究対象から直接得たデータ（「一次データ」などと呼ばれる）と、公的機関が発表している統計資料の利用などにより研究対象から間接的に得たデータ（こちらは「二次データ」）に大きく分けられます。いずれにせよ、私たちは、論文などに掲載されたデータを参考にすることができます。しかし、この段階はまだ「データ収集」の入り口にすぎません。

「データ収集」で目指すのはいわゆる「生（なま）データ」への接近です。生データとは、「加工」されていないデータのことで（ここでいう加工とは、自分の都合のよいデータに変えてしまう「改ざん」とはまったく異なります！）。わかりやすく論文などに提示されているデータのほとんどは、著者の方針にしたがって整理や分析が加えられています。たとえば、単位を揃えたり、著者が設定した区域の森林面積を合算したり、年次を区切って表したり、あるいは一定期間の減少率を算出したりすることも、簡単ですがデータの加工といえるでしょう。そうすることで、著者は自分の主張を支持する根拠をわかりやすく提示し、読者の説得を試みているのです。

加工が施される前のデータ—相対的に生データに近いもの—を入手し、読み解くことで、自分なりの判断を下すことができます。論文などでは掲載データの出所（source）を明示していますので、この情報から加工前のデータを目指します。公開されている二次データが出所であれば、比較的容易に得られます。たとえば、「東南アジアにおける森林消失速度」を示した表の出所に、ある国際機関の年鑑（たとえば『FAO Yearbook』）があげられていたならば、前述のオンラインの蔵書目録から所在が確認できます。また、ウェブサイト（たとえば「FAO Stat」）などでデータを公開している機関も少なくありません。これらの情報ソースから、表の数値を算出する元となった実数のデータ（ここでは各国各年の森林面積）が得られると同時に、関連する周辺データも広範に参照できます。こうして得られた加工前のデータを自分なりに検討してみると、論文の著者とは異なる見解に至ることもあるでしょう。最新のデータから近年の新しい傾向が読みとれるかもしれませんし、既存の研究が注目してこなかった事実に気づくかもしれません。さらには、これらの作業を通じて、より踏み込んだ新たな疑問や関心が見つかることもあるでしょう。

もう一歩すすんだデータ収集のためには、二次データであっても距離と言語の壁を越えていきます。たとえば、統計データであれば、国際機関だと主に国レベルのデータを扱っていますので、地方行政区レベルの

数値を得るためには、当該国が公表している統計資料が必要になります。対象が先進国だとインターネット経由で手に入れられる資料も多くあり（たとえば日本国農林水産省ウェブサイトなど）、新興国・途上国であっても一部のデータをウェブ上に英語で公開している国もあります（ベトナム統計総局 General Statistics Office of Vietnam ウェブサイトなど）。しかし、やはり多くの国において、現地語のみで作成され、当該国（あるいは当該地区）でしか入手できない資料が大半です。そこで、現地での資料探索や翻訳作業が求められます。とはいえ、思い通りのデータはなかなか得られませんから、完全なデータ収集にはあまり拘泥せず、手元のデータを最大限活用する姿勢も大切です。（主要な資料であれば、過去に実施された研究プロジェクトの成果物として英語や日本語で公開されていることもあります。たとえば京都大学東南アジア地域研究研究所「東南アジア逐次刊行物総合目録データベース」など。）

さらに、取り組んでいるテーマにもよりますが、自分で現地調査をしないと一つまり一次データを収集しないと一満足できない人もいるでしょう。たとえば、地域住民や行政官への聞き取り調査（インタビュー）や参与観察、劣化した林地への踏査などを通じたデータ収集です。場合によっては何かを測定したり、サンプリングしたりすることもあります。国内外限らずフィールドワークによって納得できる十分なデータを得るのは容易ではありません。しかし、仮にデータ収集としては十分な成果が得られなかったとしても、現地に臨む経験は貴重です。調査前には無味乾燥なものに感じられた統計数値データから自分なりのイメージが描けるようになったとすれば、それも大きな収穫です。将来の学修や研究の指針にも少なからず影響を与えるでしょう。

4. 定量的データと定性的データ

ここまで、「森林面積」など、数値として表すことができるデータ（「定量的データ」などと呼ばれる）を例としてあげてきましたが、数値化して扱うことに馴染まない情報（こちらは「定性的データ」）にも少し言及しておきます。たとえば、ある国の森林保全政策の内容や変遷、ある地域の住民が森林から得る産物の種類や利用方法に関する情報などがこちらに含まれます。地図情報や位置情報、写真や動画などの画像情報なども、この定性的データの範疇に入るでしょう。これらのほとんどは記述によって表されるため、データを扱っているという感覚は弱いかもしれません。しかし、定性的データの収集に臨む際にも、定量的データと同様に、身近にアクセスできる二次データのその先には無数のデータ（とその収集方法）があることを頭の隅においておけば、学修・研究段階に応じたデータ収集計画がたてられます。

5. 文献・データの探し方

(1) 「ある」と思って探す！

研究論文や書籍などの文献はデータ（二次データ）の宝庫です。私たちがインターネットや大学図書館を通じてアクセスできる文献は膨大な数にのぼります。その中から求めるものを効率よく見つけ出すため、前掲の学術情報の検索サービス「Google Scholar」や「CiNii Articles」などが利用できます。これらをうまく活用するには「検索ワード」の適切な選択がとても重要なカギとなります。最初は一般的な語句を組み合わせることで広い範囲を検索しながら、できるだけ多くのヒットした文献のタイトルや要旨を確認していくのがよいでしょう。ひとつでも自分の関心に近い論文や書籍が得られたならば、その文献内で用いられている専門用語やキーワードを次の検索ワードとして活用し、検索結果を絞り込んでいくことができます。

具体例をあげます。ある学生がテーマ「アニメによる観光開発」に関連する日本語文献を立命館大学図書館ウェブページから CiNii を使って調べようとしています（表）。はじめに検索ワードを「観光開発」ひとつで検索してみたら 1044 件の文献情報がヒットしました。タイトルだけを見ていくにもこれは多すぎると考え、

「観光開発」に「アニメ」を加えて検索してみました。すると結果は1件だけ。「アニメ」の代わりに「聖地巡礼」にしてみたら0件——。もしかすると、この時点で「アニメによる観光開発」についての（日本語で発表された）先行研究はほとんどない、と考えてしまう人もいるかもしれません。しかし、「観光開発」から「開発」を削除して、より一般的な用語「観光」を検索ワードにすると検索結果は大きく変わります。「観光」＋「アニメ」だと170件、「観光」＋「聖地巡礼」だと107件がヒットします。この数ならばひとつひとつ見ていく意欲も湧いてくるはずです。さらに、検索結果に示された論文のタイトルをみていくと「コンテンツツーリズム」という専門用語が論文タイトルに多用されていることに気づくかも知れません。これを検索ワードに使える文献をより効率的にリストアップすることができるでしょう。

表. 「CiNii Articles」による論文検索の結果件数

検索ワード	結果件数*
「観光開発」	1,044件
「観光開発 アニメ」	1件
「観光開発 聖地巡礼」	0件
「観光 アニメ」	170件
「観光 聖地巡礼」	107件
「コンテンツツーリズム」	245件
「コンテンツツーリズム アニメ」	76件
「コンテンツツーリズム アニメ 国際」	10件

*: 2021年1月29日の検索結果

また、各論文の最後に掲載されている参考文献／引用文献／リファレンスのリストは、その論文が発表される「以前」を探索する上で貴重な道標となり得ますし、Google Scholarの「引用元／Cited by」リンクから論文公表「以後」の新しい関連文献を探す際に有効です。

「関連文献やデータはどこかに必ずある」という心構えで臨むことが肝心です。見つからないときも、（どこかにあるけど）「まだ」見つかっていないだけなのです。検索ワードが多すぎないか、限定的すぎないか（別の一般的な語句に替えてみる）などを再検討してコツコツ探してみましょう。また、探していた文献が大学図書館の蔵書にもあるときは、開架書架や書庫へ足を運んでみましょう。書庫は分野別に分類されているので、目的の文献の周囲の棚を眺めてみると有用な資料に思いがけず出会えることもあります。

（2）「ない」ときはチャンス！？

どうしても自分の望んだイメージ通りのテーマを扱った文献が見つからないときは、その「イメージ」を少し広げる努力をするとよいでしょう。たとえば、「インドネシアのストリートチルドレン」の研究をすすめたけれどぴったりの文献が見つからない場合などは、インドネシアにこだわらずフィリピンやタイ、ベトナムのストリートチルドレンを扱った文献なども、大いに活用できる先行研究となるはずです。また、「ストリートチルドレン」というワードを含まずとも、インドネシアの「初等教育問題」や「都市の貧困問題」を扱った文献に研究をすすめるためのヒントや有用なデータが潜んでいるかもしれません。文献収集で困ったときには、このように文献探索の範囲を拡げてみましょう。先に文献やデータは「あると思って探す」ことが大事だと言いましたが、いつも都合よくテーマに完全に合致した情報が得られるとは限りません。研究がすすみテーマや対象が絞り込まれていけば、その傾向はより強まるでしょう。それはあなたが先駆者になる

可能性を示してもいます。

とはいえ、欲しいデータが得られないときはどうすればよいのでしょうか？ また、うまく得られた場合でも、データが断片的で（自分の要望に対して）完全には応えてくれないケースも多くあるでしょう。たとえば途上国の統計データは一般には公開されていなかったり、信頼性に問題があったり（異なる値のデータがあるなど）、そもそもデータ自体が存在しなかったりという課題にはしばしば直面します。努力して探しても見つからないときには、得られるデータを最大限に活かす工夫を試みます。たとえば、D国の木材「輸出」量のデータが直接的には得られないときには、D国の貿易相手国の「輸入」統計をしらみつぶしにチェックすることで推定量が算出できるかもしれません。E国政府が公表するケシ（麻薬）栽培面積値と国際機関やNGOの推定値に大きな乖離があるときには、その差異の背景を考察することこそ研究上の意義を見出せられるかもしれません。データが得られないときこそ、あなたの腕の見せどころ、独自のアイデアを発揮する好機なのです。

6. おわりに

疑問や関心の明確化、身近な二次データの収集（生データへの接近）、さらに踏み込んだデータの収集、といったデータ収集のいくつかの側面を大まかに紹介してきました。これらは順々にひとつひとつ完了させていくような性質のものではありません。ことさら意識せずとも、同時並行してすすみ、それぞれの段階が相互に影響を与え合うことでしょう。外から見たら順を追ってすすめたように見える研究成果でも、実は現地調査で得たデータや経験からスタートした（あるいは途中で大きく方針転換した）ものは少なくありません。そう考えると、先の「2. データ収集以前」では否定しましたが、「熱帯林保全の現状に興味をもったから」「いきなり東南アジアや南米へ飛び立つ」のも、ひとつのやり方なのです。

最後に、「1. はじめに」であげた戸惑いを思い出してみてください。担当教員の言う「テキストの熟読」とは、テキストだけを読めばよいのではなかったのです（もちろん内容の正確な理解は大前提です）。そして、「内容を批判的に掘り下げる」とは、ある主張の根拠として示された事実まで調べあげて自分でも考えてみることであり、「発展的に深める」とは、新たな問いを見つけてそれに取り組むこととも言い換えられるでしょう。基礎演習の枠内ではさまざまな制約があるでしょうが、データ収集に地道に取り組み「掘り下げて」「深めた」報告にはきっとクラスの仲間が活発な議論で応えてくれます。そして何より、あなた自身の知的好奇心が満たされ、発表にも力がいはいに違いありません。どうでしょう、基礎演習の報告順がまわってくるのが待ち遠しくありませんか？

7. 参考となる文献

- 1) ハンス・ロスリング、オーラ・ロスリング、アンナ・ロスリング・ロンランド（著）上杉周作、関美和（訳）（2019年）『FACTFULNESS（ファクトフルネス）10の思い込みを乗り越え、データを基に世界を正しく見る習慣』日経BP社

■公衆衛生の専門家でもある著者の意図はサブタイトルに込められている。具体的トピックとして貧困問題を扱っているので国際関係学部の学生にとって読み進めやすいはず。

- 2) ダレル・ハフ（著）高木秀玄（訳）（1968年）『統計でウソをつく法』講談社

■研究者やマスコミの主張を鵜呑みにせず、根拠として示されたデータをじっくり検討する姿勢は大切です。この本では、悪例を通じて、データの取り方・読み方・使い方を考える。古い本ですが、読むとデータと向き合う意欲が出てくる。

- 3) 椎野若菜、白石壮一郎（編）（2014年）『フィールドに入る』古今書院

■2014年より刊行中の「100万人のフィールドワーカーシリーズ」(全15巻)の第1巻。シリーズを通じてフィールドワークにまつわる、あらゆるトピックを新進気鋭のフィールドワーカー達が語る。

- 4) 柳澤雅之 (2019年) 『景観から風土と文化を読み解く』 京都大学学術出版

■こちらもシリーズ本のひとつ。フィールドで見聞きする様々な種類のデータから地域像を描くための手法を論じる。京都大学東南アジア地域研究研究所の地域情報学プロジェクトの成果。

- 5) 京都大学大学院アジア・アフリカ地域研究研究科、京都大学東南アジア研究所 (編) (2006年) 『京大式フィールドワーク入門』 NTT出版

■さまざまな学術論文を事例として取り上げ、試行錯誤の過程を具体的に紹介することで、フィールドワークによる研究手法(データの収集や分析も含む)を解説している。読むと現地調査に出かけたくなる。

- 6) 和田信明、中田豊一 (2010年) 『途上国の人々との話し方—国際協力メタファシリテーションの手法』 みずのは出版

■国際協力の現場で培われたフィールドワーク技法を理論化する試み。タイトルに違和感をもつ人もいるかもしれないが、研究志向のフィールドワークに通じるところもある。

- 7) 宮本常一 (1984年) 『忘れられた日本人』 岩波文庫

- 8) 佐野眞一 (1996年) 『旅する巨人—宮本常一と渋沢敬三』 文芸春秋

■前者は、偉大な民俗学者の代表作。「定性的データ」から紡ぎだされた珠玉の作品ともいえるか。後者は、彼の足跡を追ったノンフィクション。2冊あわせて読むと、圧倒されるが、人の話を聞きに旅に出たくなる。

- 9) 沢木耕太郎 (1987年) 『紙のライオン—路上の視野 I』 文春文庫

■多様な著作をもつノンフィクション・ライターが取材や表現方法について述べた本。他の著作『テロルの決算』や『一瞬の夏』などから、興味のあるテーマを扱ったものを選んで、あわせて読むとよりおもしろい。

執筆者：松田 正彦

執筆日(更新日)：2007年7月30日 (2021年2月1日)