

## 第5章 「社会学文献情報データベース」の現状と課題

周藤 真也

(早稲田大学社会科学総合学術院)

### 1. はじめに

日本社会学会データベース委員会において作成している「社会学文献情報データベース」は、構想から 10 年、本格的な運用からも数年以上が経過し、その間、データベースを取り巻く環境は大きく変化してきている。

本稿では、「社会学文献情報データベース」の現状を分析し、課題をまとめる。

### 2. データ収録状況

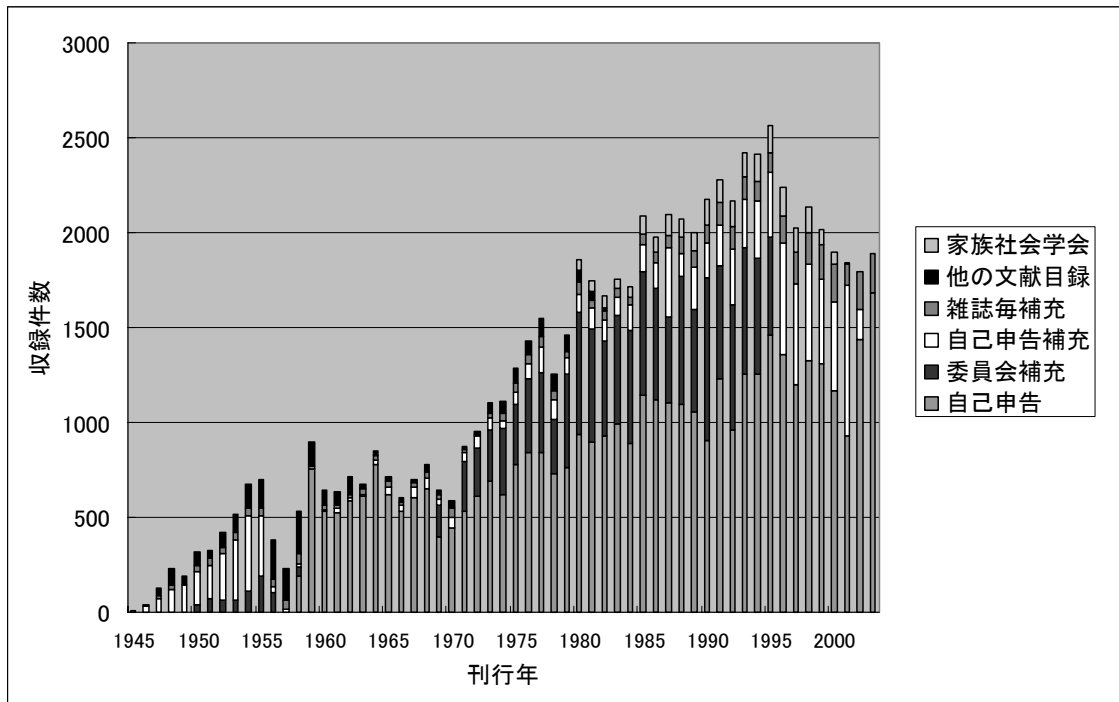
「社会学文献情報データベース」は、①日本社会学会の会員による毎年の研究業績の自己申告（1958 年～）および収録漏れの自己申告、②社会学評論編集委員会による毎年の文献目録の補充（1969～1995 年）、③日本家族社会学会による「家族社会学文献目録」および同学会の会員による毎年の研究業績の自己申告、④日本社会学会が作成した各種文献目録、⑤日本社会学会の機関誌収録文献、を中心に文献情報を収録し、形成されたものである。これらデータソースの詳細については、巻末の資料『社会学文献情報データベース』データソース一覧」にまとめてある。

図1は、最新バージョンのデータ（Version 7.1、74,923 件）をもとに、1945 年以降の各年に刊行され、データベースに収録されている文献情報の件数をデータソース別にグラフにしたものである。

会員の自己申告制にもとづくデータベースの構築は、当然ながら、学会員ひとりひとりの協力に成り立っている。図1から明らかのように、刊行年別の収録件数は、1970 年代に 1,000 件を超え、1980 年には 1,500 件超、1980 年代中頃からは 2,000 件前後を推移している（最大は、1995 年の 2,567 件）。そのうち、毎年 1、2 月頃に行われる社会学文献調査において収集された前年刊行の文献情報は、1980 年代より 1,000 件前後を推移している。

しかしながら、現在のこうしたこと文献調査による文献情報の収集状況は、2004 年末の会員数が約 3,500 名であることを考えると、決して収録率が高いとは言えないと思われる。会員 3,500 名のうちには、既に現役を退かれた方や、まだ研究業績のない若手の院生、付き合いや様々な事情から会員ではあるが社会学関連の研究論文は基本的に執筆しない近接領域の研究者（社会学を主たる研究領域とはしない研究者）、などがある程度（あるいは相当数）含まれていることは予想されるが、1 会員あたりの毎年の収録件数の平均が 1 件にも満たない現状は、実際に会員が発表している文献数に対するデータベースへの収録率はかなり低いことが予想される（例えば、会員 3,500 名、1 会員平均年 2 件で、データベー

図 1：刊行年別収録件数



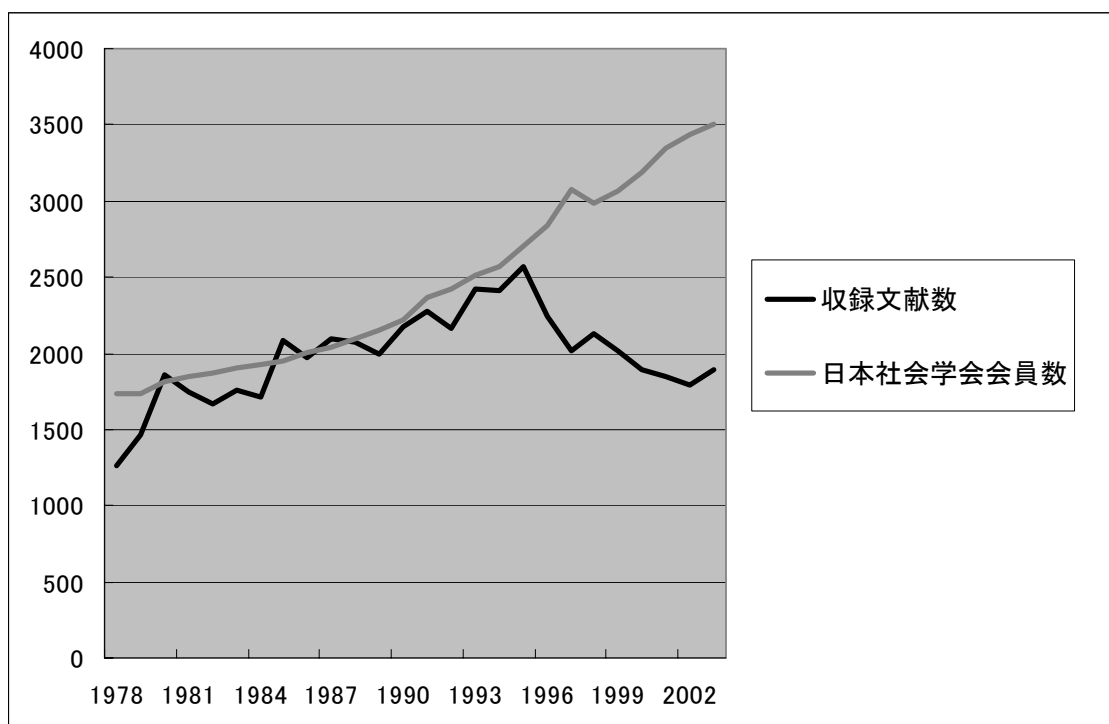
〈解説および付記〉

- ・このグラフは、データバージョン **Version 7.1 (74,923 件)** に、脱落している 1977 年「社会学文献目録」のデータ件数 **841 件** を加えて作成してある。
- ・会員の自己申告による文献目録の作成がはじまったのは、**1958 年**刊行のデータからである。
- ・**1956 年**に会員の自己申告による **1945 年～1955 年**分の文献目録作成が行われている（上記グラフでは、「自己申告補充」として扱った）。
- ・**1958 年**以前には、社会学評論編集委員会の編集による文献目録が『社会学評論』に収録されている（上記のグラフでは「委員会補充」として扱った）。
- ・これら文献目録の作成方法の変更の過渡期にあった、**1956 年**後半から **1957 年**にかけては、文献目録の作成が行われていない。**1956,57 年**刊行の収録件数が少ないのは、そのためであると考えられる。
- ・**1959 年**刊行のデータは、会員の自己申告に加えて小山隆氏が補充を行っているが、区別がつかないため「自己申告」分に含めた。
- ・少なくとも **1964 年**より国立国会図書館等での資料をもとに文献目録の充実が図られるようになったが、これも区別がつかないためすべて「自己申告」分として取り扱っている。
- ・国立国会図書館の資料による文献目録の補充が、独立して『社会学評論』に掲載されるようになったのは、**1969 年**刊行分が最初である。**1970 年**刊行分の補充はなく（**1969 年**刊行分が掲載された『社会学評論』**22 卷 1 号**の杉之原寿一氏による「編集後記」に、「従来のように国会図書館員の協力をうることが困難となった」とあり、「文献目録補充」の作成をめぐる過渡期であったことが判る）、**1971 年**刊行分で復活し、以後 **1995 年**刊行分まで続いた。
- ・データベース作成を前提として社会学文献調査を行うようになったのは、**1996 年**刊行分からである。この年より、「文献目録補充」が廃止されたこともあり、刊行年別の収録件数はその前年がピークとなっている。
- ・従来からの「調査用紙」の提出に加えて、**Web 入力**による文献情報の提出が始まったのは、**2002 年**刊行分の文献調査からである。
- ・**2003 年**刊行分の文献調査は、「既登録文献データ確認調査」と同時に行われている。

スへの収録が年 2,000 件とすると収録率は、28.625%ということになる)。実際に、データを見てみると、こまめに文献情報を登録している熱心な会員がいる一方で、学界等で活躍されているにもかかわらず、近年の文献情報が登録されていない会員も散見される。

図2は、近年の日本社会学会の会員数の伸びと、データベース収録文献数を比較したものである。1995年までは、会員1人あたりの平均にすると毎年ほぼ1件の文献情報が収録されているが、「文献目録補充」が廃止されたことを差し引くとしても、1996年以降の会員数の伸びに対して、収録される文献数は横ばいを示しており、近年の日本社会学会大会等における報告件数の増加や、論文を発表可能な媒体（雑誌等）の増加ということを考えても、社会学関連の文献の総数が、減っているという事態は考えにくく、収録率が相対的に低下していることが予想される。

図2：会員数の伸びと収録文献数



※会員数データは、毎年の大会開催時のもの（日本社会学会事務局による）。

### 3. データベースを取り巻く環境の変化

前節において確認したように、「社会学文献情報データベース」は、データベース化以後、収録率が相対的に低下していることが予想される。これは、単なる収録率の低下であるだけでなく、日本の社会学研究者および社会学教育における「社会学文献情報データベース」そのものの地位低下を伴っていることが考えられる。このことを考えるとき、「社会学文献情報データベース」の構築が始まった、1990年代後半以降の、「社会学文献情報データベース」を取り巻く環境の変化を無視することはできない。

本節では、「社会学文献情報データベース」の構築開始以後の環境の変化を、1) 重複する領域のデータベースの整備、2) 書誌情報データベースから全文情報の電子化への重点の移動、3) 電子情報をめぐる経験の位相の変容、の3点から分析していく。

#### 1) 重複する領域のデータベースの整備

近年、「社会学文献情報データベース」の存在意義が低下しているとするならば、日本において社会学の研究・教育において、「社会学文献情報データベース」が果たすべき役割・機能を代替するものが充実してきているということが考えられる。この点について、以下にまとめる。

##### ①雑誌目次データベース

2005年4月現在の「社会学文献情報データベース」のVersion 7.1において、雑誌論文は、42,612レコードあり、これは、全レコード71,959件中の56.9%を占める。このように、日本の社会学者の発表する社会学関連文献のうち、雑誌論文は中心的な研究成果の発表手段となっている。これら雑誌論文が掲載される雑誌は、社会学およびその近接領域の学会等の団体が発行する機関誌、各研究者が所属する大学・研究機関の発行する研究紀要が多くを占めている。

しかしながら、これら雑誌収録文献の情報は、日本においても、学術情報センター（→2000年4月国立情報学研究所→2004年4月大学共同利用機関法人 情報・システム研究機構 国立情報学研究所）において、学術雑誌目次速報データベース（＝研究紀要ポータル→2004年9月学術コンテンツ登録システム）（2005年4月論文情報ナビゲータ（CiNii）<http://ci.nii.ac.jp/> に収録）において、主として紀要に掲載された文献の書誌情報が、各大学の附属図書館などの協力により整備されてきた。また、国立国会図書館においても、「雑誌記事索引」が整備され、NDL-OPAC <http://opac.ndl.go.jp/> のひとつとして Web 上での検索が可能になった。これらのデータベースは、現状ではまだ網羅性に欠ける側面もあるが、日本の社会学者が学術雑誌・研究紀要等に発表する論文等の文献の何割かは、これらのデータベースにより検索可能となっている。

##### ②書籍データベース

日本の学術研究者が利用する書籍データベースとしては、国立情報学研究所総合目録データベース（WWW 検索サービス NACSIS-Webcat <http://webcat.nii.ac.jp/>）がよく知られている。各大学図書館の連携のもとに作成しているデータベースであるため、研究書の網羅性が高く、洋書も含まれており、当然のことながら社会学関連の研究書の書誌情報もこれに収録されるため、社会学関連の書籍を検索する際にも、敢えて専門データベースである「社会学文献情報データベース」を利用することなく、NACSIS-Webcat が用いられることが多くなってきているように思われる。特に、各大学の附属図書館における蔵書管理のオンライン化が進み、また蔵書検索も従来のカードや目録による検索から、電子化（OPAC）の利用が一般化したため、1990年代後半には、各大学図書館においてオンライン化以前の蔵書カードや蔵書目録の電子化もすすみ、一層書籍データベースが充実するこ

ととなった。

また、日本国内で出版された書籍の網羅性においては、国立国会図書館の JAPAN/MARC データベースについても触れなければならない。国立国会図書館は、日本国内で出版されたすべての出版物を納本することになっていることから、網羅的に出版物が収集・保存され、「日本全国書誌」の作成なども行われている。現在は、国立国会図書館の蔵書検索が Web 上で行うことができるようになってきているが（国立国会図書館蔵書検索システム NDL-OPAC <http://opac.ndl.go.jp/>）、早くからデータベース化の進んでいた JAPAN/MARC が書籍検索において果たした役割は大きかった。

また、編書に収録された各論文や、分担執筆による書籍の目次情報は、従来はオンラインでの入手が困難であり、「社会学文献情報データベース」のオリジナルな特性のひとつであったのだが、これも㈱トーハン、日本出版販売㈱、㈱紀伊國屋書店、日外アソシエーツ㈱の4社による共同構築による Book データベースが、これらの情報も一部サポートしており、インターネット上の書籍販売業者のデータベースと連携されることを通して、一般にも情報が入手しやすくなった。これら商業サイトは、そのほかにも JAPAN/MARC データベースとの連携や、各出版者から提供される情報と連携させることを通して、書籍の概要などを含め、かなりの情報が得られるようになってきている。

### ③研究業績データベース

「社会学文献情報データベース」は、日本社会学会の会員による業績目録データベースとしての機能も持っている。しかしながら、これについても、各大学・研究機関等の所属機関に提出する業績リストのデータベース化が進んだり、科学技術振興事業団（JST）による「研究開発支援総合ディレクトリ」（ReaD <http://read.jst.go.jp/>）が整備され、国立情報学研究所による「研究者ディレクトリ」が統合されたりするなど、同様な機能をもったデータベースの整備が進むこととなった。

これらのことから、「社会学文献情報データベース」は、重複する内容をもつデータベースの整備が進むにつれて、データベースの構想時や運用開始時と比較して、相対的にその価値と役割が低下してきていると考えられる。

## 2) 書誌情報データベースから全文情報の電子化への重点の移動

先述したように、この 10 年のあいだに、書誌情報データベースが充実するにつれて、関連する情報のデータベース化とともに、新たな段階として、全文情報の電子化についての取り組みが進みつつある。

たしかに、全文情報の電子化は、著作権・版權の問題や、書籍・雑誌が販売物であることとの関係、またインターネットを通じた公開は、ユーザーの範囲を制限することが難しいなどの問題から、その歩みは遅いが、もともと無償で各大学図書館等に配布している研究紀要は、まず先に電子化の取り組みがなされるようになってきている。たとえば、社会学関連で言えば、『関西学院大学社会学部紀要』は、全号が電子化され、インターネット上で試験公開されている（<http://www-soc.kwansei.ac.jp/kiyou/index.html>）。また国立情

報学研究所の「電子図書館サービス (NACSIS-ELS)」(現在は、「NII 論文情報ナビゲータ (CiNii)」<http://ci.nii.ac.jp/> 内において収集された情報が公開されている)が無償で学会誌等の電子化を行うサービスを開始したため、これを利用して、学会誌を電子化する学会も出てきており、社会学関連の学会においても、日本犯罪社会学会の『犯罪社会学研究』や、日本教育社会学会の『教育社会学研究』は、このサービスを利用して電子化されており、日本社会学会の『社会学評論』についても電子化の動きがある。

たしかに、現在の電子化された全文情報は、画像情報として提供されることが多く、使い勝手が悪いといった問題点はあるが、このように、書誌情報データベースから、全文情報の電子化へと重点の移動は、(独立行政法人 日本学術振興会) 科学研究費においても、研究公開促進費(データベース)において、平成 17 年度から「学術誌データベース」という項目が新設され(これにより、旧来からのデータベースは、「研究データベース」という項目で呼ばれる)促進されるようになってきている。

なお、日本社会学会の発行する英文雑誌『IJJS (International Journal of Japanese Sociology)』も、Blackwell 社で発行するようになった 10 号(2001 年)以降は、電子ジャーナルとしても有償で提供されており、こうした有償のものはそれぞれ別個のシステムにおいて提供されるといった問題点はあるが、いずれにしても電子化の焦点が全文情報に移動していることには違いない。全文情報の電子化にあたっては、当然ながら目次情報も電子化され、データベース化されることから、これらは、書誌情報データベースを代替する機能をもっている。

### 3) 電子情報をめぐる経験の位相の変容

これらのことを通して、「社会学文献情報データベース」は、近年の電子情報をめぐる環境の変化に否応なくさらされ、その位置づけが変わりつつある。

このことに触れるとき、1990 年代後半に起こったインターネット利用の普及は、電子情報に対するわれわれの経験の位相を大きく変容させることとなったことを無視することはできない。それまで電子化されデータベース化された情報は、CD-ROM 等の媒体により、専用のソフトを使用し、限られた範囲内で提供されることが主であった。しかし、インターネット利用の普及は、WWW (World Wide Web) に代表されるように、こうしたスタンドアロンあるいはイントラネット的な電子情報の利用から、ネットワークを通して、広く(かつ多くは無償で)提供されたものを利用することの経験を一般化させたのである。

こうしたことから、「社会学文献情報データベース」においても、本格的な運用をはじめた 1998 年から学術情報センター(→国立情報学研究所)情報検索サービス(NACSIS-IR)による有償の公開(2005 年 4 月より、NACSIS-IR の後継となる「学術研究データベース・リポジトリ」<http://dbr.nii.ac.jp/>により無償で公開)だけでなく、データベース委員会独自のサイトにおいて、データベースを公開してきたことは意義深いことではあった。しかしながら、WWW それ自体が、巨大なデータベースに喩えられるように、あるいは、WWW の概念、あるいはその基礎にある HyperText の理念に示されるように、WWW で公開される情報は、それ自体が、単独で存在するものというよりも、インターネット上の資源の一つとして、われわれに経験されるようになってきている。そうした電子情報をめぐるわ

れわれの経験の位相の変容は、単独の文献目録や文献データベースというものの地位を相対的に低下させてしまったと考えられるのである。

#### 4. 現状のデータベースの問題点

以上のように、「社会学文献情報データベース」の置かれている状況は、構想以来、大きく変化してきていると考えられる。本節では、それによって明かになってきている「データベース自体の問題点」を以下に整理してみよう。その際には、電子情報化に伴う「データベース化がもたらしたもの」に注意を払う必要がある。

##### 1) データ収録率とデータ検索経験

先に述べたように、「社会学文献情報データベース」に収録される文献情報の収録率は高くはなく、また近年低下してきていることが予想される。これは、本データベースは、日本社会学会会員の「自己申告」による文献情報の収集を中心として構築してきたことに付随する問題であるのだが、しかしながら前節で述べたデータベースをとりまく環境や、われわれの経験の位相とも密接にむすびつき、データベース化を行ったことによる固有の問題との相乗効果が起きていることが考えられる。

「社会学文献情報データベース」のもととなった『社会学評論』に掲載されていた「社会学文献目録」という媒体は、日本社会学会が定めている研究領域別に掲載されるということもあって、目録に掲載されることにより、その分野で研究業績を挙げていることを、同じ領域の研究者や、学会員全体に知らしめる効果をもっていた。しかしながら、データベース化は、目的をもって検索しない限り、われわれの眼前にその結果は現れてこないため、どのようなデータが収録されているのかについて不可視性を帯びてしまう。それとは反対に、ひとたび目的をもって検索すると、目的とするデータが収録されているか否かが一目瞭然となってしまふ。例えば、よく知っている研究者の名前や、編集・刊行などに携わっている社会学関連の雑誌名で索いてみるならば、どのくらい収録されているかは、一目瞭然となる。その領域を専門とする研究者であれば、主要な文献の多くは既知であるため、検索をすれば、どのくらい収録されていないかということが、体感されるであろう。

「文献目録」においては、『社会学評論』の掲載号のページを繰って探さなければならず、載っているかどうかを確かめるには、それなりの労力を要するし、また載っていないとしても、気づきにくいのであるのだが、データベース化はそうしたことがらで可視化されるのである。

こうしたことから、データベースは、収録率が低い(=使えない)、とひとたび判断されたならば、ますます利用率は低下し、また、データベースを構築する意欲(すなわち会員が自らの発表した文献を登録する意欲)も損なわれていく可能性が考えられる。前項で述べたように、「社会学文献情報データベース」の収録率が、芳しくないとするならば、こうした可能性は増大していると考えざるを得ない。

## 2) 作成労力の増加

以上のことから、「社会学文献情報データベース」が、一定レベル程度以上の「使える」データベースであるためには、「文献目録」時代と比較して、一層データの質の高さを求められるし、データの網羅性に関しても一定水準以上のものが求められるということになるといえよう。

しかしながら、それを実現するためには、それ相応の労力が必要となる。「文献目録」では、文献情報の単なる記述にすぎなかったものが、データベースにおいては、(これは電子情報の特性であるのだが) あいまいな記述や情報を許さないもので、項目ごとにきちんと整理した情報として記述される必要がある。それには、データベースの構成や、記述の形式をきちんと確定させる必要があるし、付加的な情報の記述についてのルールなども決める必要がある。データベース化によって、従来の「文献目録」では、気づかれることがなかったデータの記述の揺れが明らかになったりするし、それをどのように統一するのかわかるという問題も出てくる。また、各項目に対応する欧文(英文)項目、著者名や標題名などのカナ記述の項目、文献種別、本文の言語などの項目が設けられ、文献情報の充実が図られたため、1件の文献情報につき入力可能な項目が34項目にもものぼることとなった。校正者にも、データベースの構成やデータ記述についての知識が必要となったりするし、作業者には、テキスト処理に関する基本的な技能も要求されるようになる。

このように、データベース化は、編集・とりまとめに当たる側(すなわちデータベース委員会)において必要とされる労力を増加させた側面があると考えられる。また、文献情報を提供する日本社会学会会員においても、データ項目数の多さ、特に欧文(英文)情報を付加することは手間がかかることであるし、また、所属研究機関が研究業績をデータベース化や、科学技術振興事業団の「研究開発支援総合ディレクトリ」への研究業績の提出など、同様の情報の提出を二度三度と要請されるような状況も起きている。雇用関係や業績評価・研究者評価と直接的に結びつく所属研究機関への業績目録の提出に比べて、学会が行っている文献情報データベースの作成への協力は、相対的に優先度の低いものとして取り扱われることがあるのではないと思われる。

こうしたことから、労力を要する割には、利用価値がそれほど大きくない、文献目録、文献情報データベースの作成は、社会学の研究教育活動において、その地位が低下してきていると考えられるのである。

次に、現行データ、および新規入力データについて、具体的な問題をまとめておくことにしよう。

## 3) 現行データの問題

現在公開中のデータ(Version 7.1、74,923件)は、「社会学文献目録」やその他の資料のデータベース化以後も、データ記述の統一や、残存するミスの修正などの手が加えられているが、依然として次のような問題があることがわかっている。



#### ① 記述形式の誤り

「社会学文献情報データベース」には、文献種別に応じて、入力してはならない項目がある。現行データには、こうした入力してはならない項目にデータが入力されていたり、あるいはそもそも文献種別が誤っていたりするデータがあることが判っており、修正が望まれる。また、データの記述方法についても、揺れがある項目や、ルール通りに記述されていないデータが散見される項目がいくつかあり、統一を図る必要がある。

#### ② データ記述の揺れとその統一

各項目のデータ記述のうち、同一となるべきデータ記述に、揺れがあることがわかっている。具体的には、「著者名」、「雑誌名」、「出版者」、「出版地」などであるが、このうち、「著者名」については、これまでの作業において、ほぼ統一が完了しているが、他の項目についても統一が望まれる。

#### ③ 欧文（英文）情報の付加

「社会学文献目録」をもとにデータベース化した文献データには、基本的に「標題名」、「掲載誌名」、「出版者」などの欧文（英文）の情報が含まれていない。これらは、2004年に行った「既登録文献データ確認調査」において補充を図ったが、今後こうした取り組みを続けるとともに、日本社会学会が作成して提出した国際交流基金『An Introductory Bibliography for Japanese Study 日本研究基本書目』（原則隔年刊行）のうち **Sociology** の分より標題の英訳および未収録文献の入力を行っていく必要があると思われる。

#### ④ その他

校正が不十分で残存しているミスの修正、情報が不十分のまま収録している文献情報の調査などの課題がある。また、修正情報は、現在「富山大学サイト」において収集活動を行っているが、いっそう修正情報の収集と修正作業の行いやすいシステム作りを図っていく必要があると考えられる。

### 4) 新規入力データの問題

#### ① Web 入力システムにより入力されたデータの問題

「社会学文献情報データベース」は、2002年度の文献調査より、Web入力を併用することになった。これによって格段にデータ作成の手間が軽減され、2002年度調査における文献情報の収集件数が例年に比べてアップするなどの効果がみられた。しかしながら、現在のものには、いくつかの問題があり、いっそうの改善が望まれる。

なお、Web入力方式の開始によって、文献情報の一年中の受け付けが比較的容易になった。このため、2004年度において会員向けのアナウンスを行い、文献情報の収集に努めることとなった。

#### ・データ漏れが起りやすい項目の存在

言語コードや翻訳コードなど、いくつかのデータの漏れが起りやすい項目が存在して

いる。これらは、入力の際に、より厳密にチェックをかけることによって、ある程度、防ぐことができるようになるだろう。

#### ・翻訳文献の問題

翻訳文献は、「著者名」項目に、原著者、訳者の順で、入力しなければならないのだが、どちらかを落とすケースが頻出している。また、原著文献の情報についても、不十分なものが散見され、必要な項目の情報を確実に収集するための工夫が望まれる。

#### ・その他データベースの構成や記述形式に関わる問題

挙げられている記述例や説明書きがきちんと読まれていないため、データベースの構成や記述形式のルールに則っていないデータの記述が散見される。記述例や説明書きをよく読むよう誘導するとともに、**Web** 入力システム側で補完するよう改善することが望まれる。

### ②調査用紙によって提出されたデータの問題

**2002** 年度の文献調査から **Web** 入力システムによる文献情報の収集を開始したが、以前からの調査用紙の提出による文献情報の収集活動も継続して行っている。新規データの中核は、**Web** 入力を中心となってきているが、会員にはインターネットを利用しない会員やコンピュータの不得手な会員もあり、調査用紙による文献情報の収集活動は今後も継続的に行っていく必要があると思われる。

しかしながら、調査用紙によって提出されたデータには、文献情報の不十分なものも多く、委員会においてかなりの補充作業が必要となっている。調査用紙による提出の場合も、会員にできるだけ必要な項目を用紙に記入してもらおうとともに、補充作業の負担を軽減するためシステム作りが望まれる。

## 5. 文献情報の網羅的な収録に向けて

### 1) 経緯

日本社会学会データベース委員会は、文部（科学）省（→日本学術振興会）科学研究費補助金において、従来より「研究公開促進費（データベース）」を取得してきており、データ入力などデータベース構築そのものにかかわる経費については、これによりまかなってきた（平成 15 年度より「重点データベース」に指定されている）。この補助金は、従来からの「社会学文献目録」および日本社会学会が作成してきた他の文献目録等の資料のデータベース化、および、毎年の「社会学文献調査」において発生する新規データの入力および校正作業の経費に充当してきたが、文献目録等の資料からの入力も主なものについてはほぼ完了しており、次のステップとして、各社会学関連雑誌に収録された文献の書誌情報データベースへの網羅的な収録に取り組むこととなった。またあわせて、他の方法における文献情報の網羅的な収集とデータベースへの収録についても、検討をはじめているところである。

## 2) 考えられる各種文献情報収集活動

文献情報の網羅的な収集とデータベースへの収録にあたって、次のような方法が検討されよう。

### ①雑誌・紀要等

社会学および社会学近接領域の学会等の研究団体が発行する機関誌・研究雑誌、および、各大学の社会学部・社会学研究室等の学術研究機関が発行する研究紀要に収録された文献の書誌情報を網羅的に収集し、未収録のものをデータベースに収録する。

### ②講座、叢書、シリーズ等

社会学関連の講座、叢書、書籍のシリーズについて、書誌情報を網羅的に収集し、未収録のものをデータベースに収録する。

### ③研究者毎の文献目録の収集

社会学研究者毎に、著書、論文等の書誌情報を網羅的に収集し、未収録のものをデータベースに収録する。大学等の機関の研究紀要には、研究者の定年退職時に、履歴などとともに、業績目録が掲載されることも多く、これらを基礎資料として収集し、活用することが考えられよう。特に、戦前期の社会学者や、戦後の社会学者でも物故者やすでに第一線を引退された研究者を中心に、文献目録の収集と調査、未収録文献のデータベースへの収録という作業に入ることが考えられる。

### ④出版者（社）毎

社会学関係の書籍を多く出している出版社毎に、社会学関係書籍ならびに収録されている各論文の書誌情報を網羅的に収録することが考えられる。協力の得られる出版社からは、あわせて書籍の概要や紹介情報の記事の収集も行い、データベースに収録することも検討されよう。

### ④海外の著名な社会学者の翻訳および関連文献

海外の著名な社会学者毎に翻訳文献の調査を行い、未収録の文献を網羅的にデータベースに収録することが考えられる。また、社会学者毎の関連文献（その社会学者の研究について書かれた文献、特に翻訳文献を中心に）についても、あわせて文献情報の収集を行い、データベースに収録することも考えられよう。

### ⑤毎年の新規データをもとに補充（編書を中心に）

毎年行っている「社会学文献調査」において収集される文献データのうち、編書ならびに編書論文については、当該編書に掲載されている他の論文が収録されていないことが多いと考えられる。これらについて、補充調査を行い、データベースに収録することが検討されよう。

#### ⑥書評対象文献

社会学関連の雑誌・紀要に掲載される書評対象書について、データベースに収録済みか否かを確認し、未収録のものについては収録することが考えられる。

#### ⑦その他各種文献リストの収集および提携

社会学関連の各領域の学会が作成している（した）文献目録や、各分野、研究テーマ毎に社会学研究者が作成した各種既存文献リストを収集し、作成団体・作成者の同意のもと、未収録文献をデータベースに収録していくことが考えられる。また、データベース作成団体・作成者との間で協力・提携関係を結び、各領域やテーマ毎に網羅的なデータ収集体制を築いていくことも考えられる。

### 3) 雑誌・紀要等の調査の実施と作業の現況

以上のように、文献情報の網羅的な収集とデータベースへの収録にあたって検討される方法をまとめてきたが、このうち、①社会学関連の雑誌・紀要等に掲載された文献、②講座、叢書、シリーズ等に収録された文献については、すでに **2003** 年度後半より作業を開始した。

社会学関連の雑誌・紀要等に掲載された文献の書誌情報の収集においては、対象となる雑誌・紀要等、数十誌程度をリストアップし、資料にある作業マニュアルに基づいて、書誌情報の書かれたページをコピーしたり、あるいは目次等のコピーに欠けている情報を補記したりすることによって、それを取りまとめるという方法によって行っている。

そのうち十誌程度については、すでに入力作業が終了し、校正段階に入っており、**2005** 年度の早い段階での公開が予定されている。

〈資料〉「雑誌・紀要等に収録された社会学関連文献の網羅的な調査」作業マニュアル

### 4) 継続的な文献情報の収集活動のために

「社会学文献情報データベース」は、これまで主として会員の「自己申告」によって文献情報の収集を行い、データベースを作成してきた。今後、それとは別に、データベース編集者側で文献情報を網羅的に収集し、データベースに収録することが加わることは、データベースの構築方法だけでなく、データベースの性格・性質を変容させる可能性をもっていることには注意を払う必要がある。

特に、雑誌・紀要に収録された文献の書誌情報のデータベースへの網羅的な収録については、今後、新しい号が出る度に、継続的に書誌データを収録していく必要がでてくる。これらのことから、社会学関連の各学会・研究団体や、各大学の社会学部・社会学研究室との連携により、継続的な情報収集とデータベースへの収録に努めていくことが検討される。

## 6. 新たなるデータベースシステムの構築に向けて

### 1) リレーショナルデータベースの構築

現行のデータにおいて、いくつかの問題があることは、前々節で述べた。なかでも、「著者名」、「掲載誌名」、「出版者」、「出版地」などの記述を統一し、それを維持していくためには、それぞれにおいて独立したデータベースを作成し、それを参照することが求められる。こうしたことを簡便に行うためには、現行のデータベースをもとに、リレーショナルデータベースを構築していくが必要になってくる。

以下、このことに関して、特に「著者名」および「掲載誌名」に関連することがらについて述べておく。

#### ① 著者（名）データベース

著者名の記述については、「著者（名）データベース」を作成し、これを参照することによって統一を図ることが検討される。現行のデータベース形式において、「著者名」項目には、漢字、カナ、英字の各項目が存在する。これらを「著者（名）データベース」に収録し、単一の著者 ID をもとに、それぞれを参照すればデータベース構成としても簡便になる。

しかしながら、著者名の記述については、その文献の現物に記載されている通りの記述でデータベースに収録する「原典主義」を考えると、過度の統一は避けるべきであろう（特に、雑誌・紀要等に掲載された文献の書誌情報の網羅的な収集においては、現物にあたるため、この問題が生じてくる）。同一著者においても、改姓、改名、「旧姓」または戸籍姓の使用および併記、筆名の使用または併用、などに対応するためには、著者 ID に対して、著者名の記述は複数あり得ることになり、別に考える必要が出てくる。しかしながら、特に著者名のローマ字表記については、ヘボン式／訓令式、長音の記述などに綴りの揺れが見られ、また著者本人による表記の統一の希望も寄せられるため、著者名記述の修正や統一に関する基準作りが必要になってくると思われる。

また、「著者（名）データベース」は、これをもとに希望者に対して著者本人に関する付加情報（連絡先やホームページへのリンク）なども収録して公開するといったことも考えられるだろう。

#### ② 掲載誌名（雑誌・紀要）データベース

掲載誌名の記述については、同様に「掲載誌名（雑誌・紀要）データベース」を作成し、これを参照することによって統一を図ることが検討される。また、掲載誌についての情報には、出版者や出版地も含まれるため、これらについても「掲載誌名（雑誌・紀要）データベース」に盛り込むか、別に「出版者データベース」、「出版地データベース」を作成し、それらを参照することが考えられるだろう。

「掲載誌名データベース」に関して、重要な課題となってくるのは、既にはじまっている雑誌・紀要毎の網羅的な掲載文献情報の収集作業との関係である。この作業においては、作業の遂行上、雑誌・紀要毎のデータ管理が必要になってくる（それらには、発行状況や

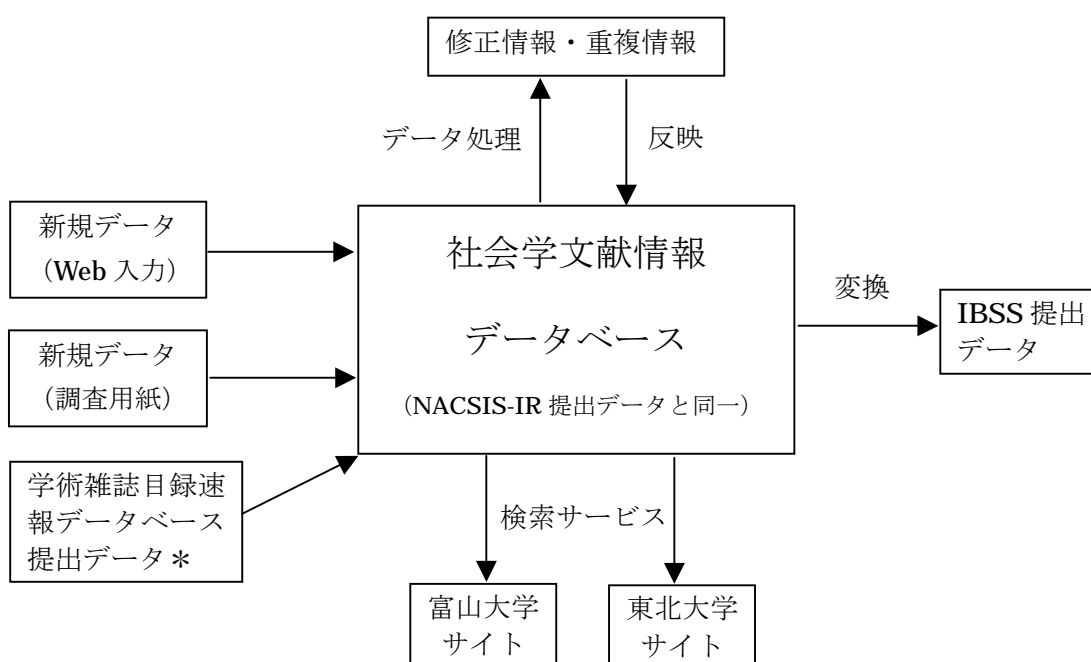
巻・号毎の発行年月や収録文献についての情報が含まれる)。こうしたことを考えると、「掲載誌名データベース」を発展させる形で、「社会学文献情報データベース」の構築のために、そのサブセットとして「雑誌・紀要データベース」を形成していくことが考えられる。

## 2) 新たなる統合的なデータベースシステムの構築に向けて

以上のように、現行データにおける「著者名」、「掲載誌名」、「出版者名」、「出版地」等の記述の統一や、網羅的な文献情報の収集およびデータベースへの収録作業は、「社会学文献情報データベース」をリレーショナルデータベースとして再構築していく作業になることが予想される。データベースの再構築に際しては、これにあわせて、現行のデータベースシステムにおける諸問題を解決し、よりデータの管理とデータベースの構築がやりやすい方式にしていくことが望まれる。

図3は、現行の「社会学文献情報データベース」の構成を図で表したものである。現行のデータベースは、タブ区切り34項目から成るテキストのNACSIS-IR提出用ファイル(2005年4月に「学術研究データベース・リポジトリ」へ移行)を基準に、これに向けた新規データの作成(東北大学サイトのWeb入力システムによるものを含む)、これをもとにした富山大学サイトでの検索システム(独自プログラムによる)および修正・重複情報収集システム、東北大学サイトの検索システム(データベースソフト「ファイルメーカー」の利用による)の構築が行われている。それぞれのシステムは、別個に動いていることは、セキュリティや作業分担の面では有利であるのだが、その反面、新規入力データや、修正情報のデータベースへの反映に時間がかかるという難点も抱えている。

図3：現行のデータベースの構成（概念図）

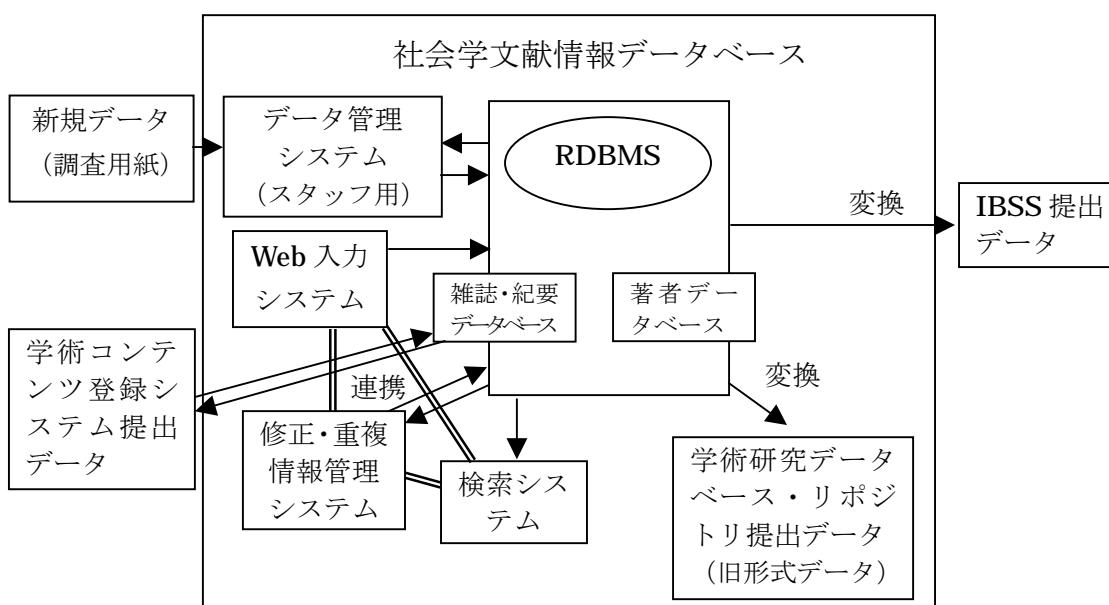


\*日本社会学会の機関誌『社会学評論』『International Journal of Japanese Sociology』の学術雑誌目録速報データベース（→「論文情報ナビゲータ」に収録）への提出データは、データベース委員会において作成しており、このデータを利用してこれらの機関誌に掲載された文献の書誌情報の「社会学文献情報データベース」への収録を行っている。なお、社会学関連学会では、北海道社会学会、数理社会学会等もそれぞれの学会内においてデータを作成して学術雑誌目録速報データベースへのデータの提出を行っている。

これに対し、図4は、リレーショナルデータベースの構築にあたり、検討される新しいデータベースシステムの構想を図により示したものである。RDBMS（Relational Database Management System リレーショナルデータベース管理システム）を核として、「雑誌・紀要データベース」や「著者データベース」と連携した「社会学文献情報データベース」を構築する。「Web 入力システム」、「修正・重複情報管理システム」、「検索システム」などデータ入力、データ公関係のシステムが有機的に接合され、学術研究データベース・リポジトリに提出される従来形式のデータは、「社会学文献情報データベース」のサブセットとして提供される。

新しい「社会学文献情報データベース」のデータベースシステムにおいては、これら全体がひとつの統合的なデータベースシステムとして構築されることが期待される。例えば、オリジナルな検索サイトにおいては、リレーショナルデータベースの特性を活かして、関連する文献データを相互に参照できるようにするなど、データ公開面においても工夫をすることが考えられるし、文献情報の収集においても、登録を促すようなシステム作りを行うことを通して、自己申告制を活かした文献データの充実に向けた工夫も考えられよう。また、付加情報、要約情報、全文情報の収録可能性など、拡張性の高い統合データベースシステムを構築することが構想されよう。

図4：新しいデータベースシステムの構想（概念図）



### 3) MySQL と PHP を用いた検索サイト構築の実験

こうした新たなる統合的なデータベースシステムの構築に向けて、RDBMS を用いたデータベースの検索サイトの構築の実験を 2005 年 1 月に実験を行った。RDBMS としてはオープンソースソフトウェアで高速性と堅牢性に定評のある MySQL <http://www.mysql.com/> を利用し、高機能 Web-DB アプリケーションの作成が容易にできることで知られているスクリプト言語 PHP: Hypertext Preprocessor <http://www.php.net/> によりプログラミングを行った。ただし、RDBMS を用いてはいるが、従来形式のデータを取り込んだだけであり、作成したデータベースおよびデータベースのテーブルは単一である。この実験にあたって、使用したワークステーションのスペックおよびソフトウェアのバージョン等の詳細は、次のとおりである。

**[Hardware]** Sun Ultra 10 / CPU: Sparc Iii 440MHz / Memory: 512MB

**[Software]** Solaris 9 + Apache 1.3.33 + PHP 4.3.10 + MySQL 4.1.9

以下、日本語で書かれた情報を含むデータベースにおいて、検索サイトを構築する際に問題になるであろうことについて、この実験において試みたことを述べておく。

日本語情報を含むデータベースにおいて、「著者名」や「刊行年」、「文献資料種別」、「本文の言語」など、検索語との完全一致が期待される場合はよいが、「標題名」などは、標題の一部やキーワードとなる日本語の単語により部分一致により検索を行うことになる。この場合、データベースシステム内部的には、ワイルドカードを含む全文検索を行うことになり、検索時間がかかったり、マシンパワーを必要としたりする。このため、たとえば検索結果が多い場合に一定の件数でページを分割する際、ページを遷移する度に再検索を行うため時間がかかるといった問題が生じたりする。本実験においては、検索結果の情報をキャッシュファイルに保存することにより、処理にかかる時間を半分以下にすることを実現した。

こうした全文検索を行う場合、あらかじめインデックスを作成することが一般的である。スペースなどによって単語の区切りが明らかな英文等の欧文の情報の場合であれば、スペースなどの区切り文字・記号で分割してインデックスを作成するシステムが、RDBMS には備えられている場合が多いが、日本語の場合はそう簡単ではなく、あらかじめ単語によるスペース区切りの「分かち書き」を行う必要がある。たとえば、形態素解析システム ChaSen (茶筌) <http://chasen.naist.jp/hiki/ChaSen/> や 漢字→かな (ローマ字) 変換プログラム KAKASI <http://kakasi.namazu.org/> などの日本語解析プログラムは、こうした「分かち書き」を行うことができるソフトウェアとして知られており、本実験においても KAKASI を使用して「標題名 (漢字)」の「分かち書き」を行い、インデックスの作成ならびに検索の実験を行った。しかしながら、日本語の「分かち書き」は一意に決められるものでもなく、検索語を「分かち書き」にして AND 検索をすることになり、検索の精度という点においては課題を残した。

作成した検索システムにおける、主な画面の見本をあげておく (図 5 ~ 8)。



図5：簡易検索画面見本



図6：詳細検索画面見本



図7：検索結果画面見本



図8：レコード詳細表示見本



## 7. おわりに

日本社会学会会員からの「自己申告」による「社会学文献目録」の作成がはじまってまもなく半世紀を迎える。また、本稿でみてきたように、「社会学文献情報データベース」の構想から 10 年以上が経過し、本データベースを取り巻く環境も大きく変わってきた。そうした意味において、戦後の日本の社会学の研究・教育活動において「社会学文献目録」あるいは「社会学文献情報データベース」どのような役割を果たしてきたのかを再考し、今後の位置づけを再検討する時期に来ているだろう。

しかしながら、「社会学文献情報データベース」には、多額の科学研究費補助金が投じられてきており、その元はいままでもなく国民の税金である。「社会学文献情報データベース」を日本の社会学界の共通の資産として、どのように活かし、どのようにその成果を還元していくのか——今まさにそのことが問われているのである。