

1995年度「社会・意識調査データベースの作成」 事業報告

新國三千代, 石井 和平, 小内 純子, 田中 一

本年度も昨年度に引き続き、文部省の科学研究費補助金「研究成果公開促進費」(データベース)を受けて「社会・意識調査データベース(SORD)の作成」事業を進めている。今年度は、新たに2名が本プロジェクトに加わりメンバは合計11名となった⁽¹⁾。また、事務局も1名増え3名となり強力になった。

今年度の主な取り組みは、素データの蓄積に向けた準備である。昨年度、本プロジェクトの各メンバが所有する素データについて、データを提供できる形に整備する試みを行った。今年度は、これをもとに素データ作成フレームワークの検討を行い、実際に素データの利用の試みを行って不備な点や問題点を洗い出す作業を行っている。

10月27日に第4回「社会・意識調査データベース」ワークショップを札幌学院大学において開催した。今回のワークショップには、政治学の分野で政治意識調査のデータベースの作成に長年関わってこられ、素データの公開を実践されている三宅一郎教授(大阪国際大学, 1996年4月からは関西大学総合情報学部)をお招きして、日本および諸外国の事例を基に素データの公開に関する問題とその対応について意見を述べて戴いた。更に、メンバの一人である直井優教授から、社会調査の素データのデータベース化および公開について問題の指摘があった。これらをめぐって活

発な意見交換が行われた。これらの議論も考慮に入れながら、ワークショップでは今後の取り組みとして次のことを確認した。

- 1) 今までに本プロジェクトが積み上げてきた成果を社会学会で発表していく。
- 2) 素データフレームワークを各委員の意見を集約して事務局側でまとめる。
- 3) 素データフレームワークを適用して、実際に素データ作成の試みを行う。
- 4) 素データの利用の試みを行い、検討すべき課題を明らかにする。
- 5) 来年度も科学研究費補助金を継続申請する。

1)については、本プロジェクトと社会学会会員との結びつきが強いことからその成果を報告する義務があることは言うまでもないが、社会調査の分野において素データをデータベース化することを研究者はどう捉えるべきなのかといった本質的な問題なども含め、社会学会の中でデータベース構築に関する議論を興していくことを意図している。2)の素データ作成フレームワークの検討では、データ提供者の負担を最小にすることを優先に考えている。公開できない項目の編集やデータを利用する際の禁止事項(例えば、被調査者や地域が特定できる分析の公表を禁じる)等も盛り込んでいる。現在のところ、素データ作成フレームワークは社会学分野の研究者のみを対象にしているが、研究者による教育における利用という観点も考慮に入れている。末尾の資料に素データ作成フレーム

NIKKUNI Michiyo, ISHII Wahei, ONAI Junko,
TANAKA Hajime
札幌学院大学社会情報学部

ワーク (SORD 方式) を添付した。3) については、本プロジェクトで実施した社会調査に関する全国調査で「素データの公開」について「可能」と答えた回答者数名 (本プロジェクトのメンバ以外) に、素データ作成フレームワークに沿って素データを作成することをお願いし、了承を得ている。これにより、素データを作成する際の問題点を明らかにしたいと考えている。4) については、委員の中で素データを利用し合い利用する上での問題を検討することになっている。

今後の事務局の取り組みとしては、「日本社会学会会員が1980年以降に実施した社会調査」の追調査を計画している。

以上をまとめると、今年度は素データの蓄積に向けた具体的な取り組みが開始された年と位置付けることができる。素データの作成では、新たにメンバ以外の研究者の協力も得ることができ、データベース活動も広がりをもってきたと言える。一方、この過程で解決していかなければならない課題も明らかになってきた。これについては、本学部の『紀要』の本号⁽²⁾で詳細にまとめているのでここでは割愛するが、社会調査の研究や教育活動さらに社会学会と連携しながら解決していくべき問題が多々あることが分かった。その意味で、札幌学院大学を中心として進められてきた本プロジェクトは、開始してから6年を経て新しい段階に入ろうとしているといっ

よいであろう。今後は、素データの作成を着実に進めていくと同時に、先に述べたように社会学会の中でデータベース化についての議論を興すという活動を行っていくことにより、継続的なデータベース化活動を可能にしていくことができると考えている。

注釈および参考文献

(1) 現在の社会・意識調査データベース作成委員会メンバの構成は次の通りである。

委員長 田中 一 (札幌学院大学社会情報学部)

委員 宇田川拓雄 (北海道教育大学函館分校)

小島 秀夫 (茨城大学教育学部)

盛山 和夫 (東京大学文学部)

直井 優 (武蔵大学人文学部)

松本 康 (名古屋大学文学部)

真鍋 一史 (関西学院大学社会学部)

*原 純輔 (東北大学文学部)

事務局 *石井 和平 (札幌学院大学社会情報学部)

小内 純子 同上

新國三千代 同上

(あいうえお順, *は新メンバ)

(2) 新國三千代: 社会・意識調査データベース (SORD) プロジェクトの成果と課題そして今後の展望 — 情報システム学の視点から —, Vol.5, No.2, (1996).

資料

「素データ作成フレームワーク」(SORD方式)

下記の要領で1)～5)を用意する。フロッピーは、DOSまたはマッキントシュでフォーマットされたものを使用する。

1) 調査票

調査に使用した調査票のコピーそのものを添付する。

機械可読化されている場合は、フロッピーにテキスト形式で格納する。

2) 素データ

- ・素データは調査票の回答結果のみからなるデータとする。
- ・公開できない項目については、作成者の判断で空白にするか記号の置き換え、番号の付け直し等を行う。
- ・素データは基本的にテキスト形式で格納するものとするが、表計算ソフトや集計用ソフト、データベースソフトで素データを作成している場合は、使用ソフト名を明らかにして、そのソフトの標準的なデータ格納形式で保存する。
- ・フロッピー数枚にも入りきらない場合は、磁気テープとする。

3) コードブック

- ・コードブックは下記の①, ②, ③のいずれかの形で作成する。
- ・コードブックはフロッピーにテキスト形式で格納する。
- ・コードブックで調査票の質問と素データの列位置の対応やデータの意味(非該当, NA, DN等)が分かるようにする。

① SPSS/SASプログラムを用いている場合

- ・データの列位置, 質問番号, 質問内容, データの内容(回答の内容, 回答にないデータ: 非該当, NA, DK等)および単純集計結果を例1の形でテキスト形式で格納する。こ

れらは、SPSS/SASプログラムで出力した結果に付加して完成する。

- ・入力プログラムと単純集計プログラムもフロッピーに格納する。

② 表計算/集計用ソフト/データベースソフトを用いている場合

質問番号, 質問文(質問見出し), データの内容(回答データ, 回答にないデータ: 非該当, NA, DK等)を例2の形でテキスト形式で格納する。質問と回答データの対応が分かるように, 画面上のレイアウトを添付する(画面印刷の結果に記入してもよい)。

③ その他の場合(独自のプログラムを利用している場合なども含む)

列位置, 質問番号, 質問文(質問見出し), データの内容(回答データ, 回答にないデータ: 非該当, NA, DK等)を例3の形でテキスト形式で格納する。

4) 素データを利用する際の条件

データ作成者は、素データを利用する際の禁止事項(例えば、「被調査者や地域が特定できるような分析結果の公表を禁じる」)や注意事項を作成し、テキスト形式で格納する。

5) フロッピーに格納されたファイル名の一覧表

提供される素データのセットは次のようになる。

- a) 調査概要情報: 事務局で調査概要情報データベースからダウンロードしたもの
- b) 調査票: 事務局で調査票を複写したもの
- c) 素データ作成者から提供された下記の3種のファイルをコピーしたフロッピー
 - ・素データ
 - ・コードブック(①の形で提供, ②, ③は事務局側で①に再編集し直す)
 - ・データを利用する際の条件

例1 (SPSS/SASプログラムを用いている場合)

SPSS/SASのプログラムの出力結果に二重下線部を付加する.

1カラム目

↓

カラム：1-2 ←カラム位置 (1桁の場合は”カラム：3”)

Q1 あなたの年齢 ←質問番号と質問文 (質問見出し)

AGE (←プログラムで表示する変数名)	Value	Frequency	Percent	Valid Percent	Cum Percent
<u>20歳未満</u>	1	124	13.7	13.7	13.7
<u>20歳代</u>	2	875	56.4	56.4	56.4
<u>30歳代</u>	3	358	25.8	25.8	25.8
<u>NA</u>	.	20	9.2	missing	
↑					
↑	Total	1377	100.0	100.0	

データの内容 (回答内容)

例2 (表計算/集計用ソフト/データベースソフトを用いている場合)

1カラム目

↓

Q1 あなたの年齢 ←質問番号と質問文 (質問見出し)

20歳未満 1

20歳代 2 | ←データの内容 (回答内容)

30歳代 3 |

NA 9 |

DK 99

例3 (その他の場合)

1カラム目

↓

カラム：1-2 ←カラム位置

Q1 あなたの年齢 ←質問番号と質問文 (質問見出し)

カラム：3

.....

以下, 上記例2と同様.