

## SSJ(Social Science Japan) データアーカイブにおける データの保存と普及

佐藤博樹

東京大学大学院情報学環教授(社会科学  
研究所(兼務))

2014年2月

### データアーカイブの社会的機能

- マイクロデータの収集、整理、保存、提供
  - データに基づく実証的研究の社会的基盤の提供
  - 同一のデータに基づいて異なる仮説による理論構築が可能:研究の深化
  - 既存データを活用して新しい研究が可能
  - 研究資源の有効活用に貢献
  - 被調査者の負担軽減、調査環境の改善
- **実証的研究における再現性の担保**

## 日本におけるマイクロデータ提供機関の類型

### ①マイクロデータを収集、整理、保存、提供の4つの機能を持つ機関：データアーカイブ

#### a)社会科学の多様な領域のデータを対象

東京大学社会科学研究所のSSJデータアーカイブ

立教大学社会情報教育研究センターの社会調査データアーカイブ

#### b)特定領域のデータを対象

レヴアアサン・データバンク(投票行動や町内会・自治会等近隣住民組織などに関するデータを収集・提供;有料)

SORD(札幌学院大学による社会・意識調査データベースプロジェクト、日本社会学会会員により実施された社会調査を収集・提供)

### ②自機関が実施した調査データを提供

家計経済研究所(家計研パネル調査)

労働政策研究・研修機構のJILPTデータアーカイブ(機構が実施した調査)

慶應義塾大学パネル調査共同研究拠点(日本家計パネル調査)

日本大学総合学術情報センター(日本大学「健康と生活に関する調査」第1回のみ公開)

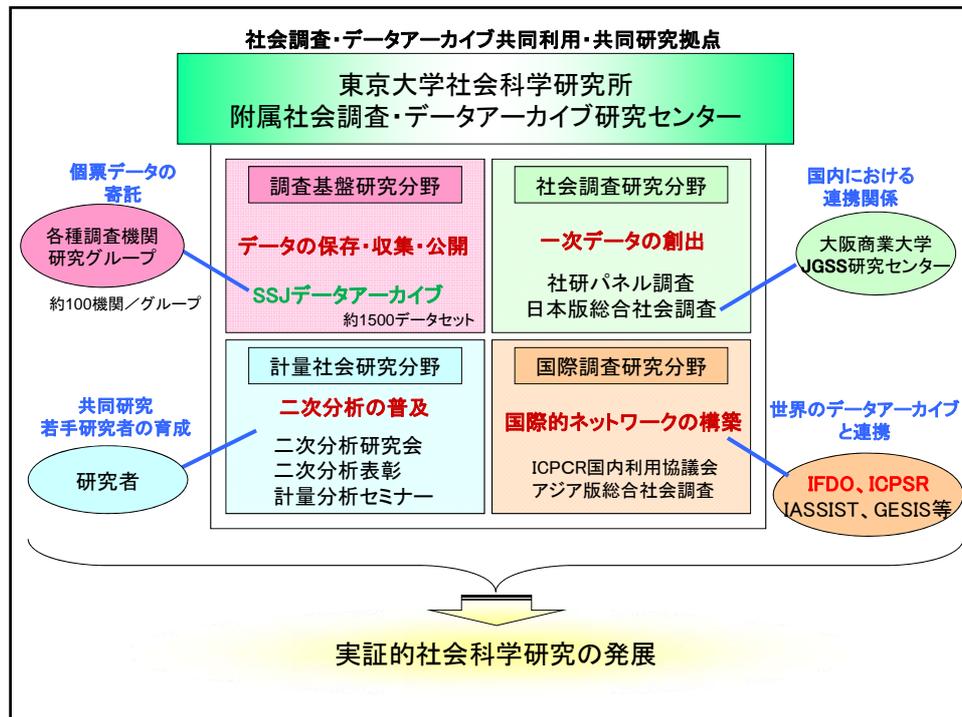
東京大学谷口研究室・朝日新聞共同調査データアーカイブ(東京大学旧蒲島郁夫・谷口将紀両研究室と朝日新聞の共同調査)

### ③政府統計の「匿名データ」の提供

独立行政法人統計センター(「匿名データ」の提供);学術研究機関等に所属する者の申請窓口としての「サテライト機関」が一橋大学、神戸大学などに設置。

## SSJ(Social Science Japan) データアーカイブの概要

- 東京大学社会科学研究所の日本社会研究情報センターに1998年に設置(2009年4月より社会調査・データアーカイブ研究センターと改組)
- 社会科学研究所で保存していた労働調査関係のデータをはじめ、民間の調査研究機関や大学等の研究者が実施した社会調査の個票データ(マイクロデータ)を収集し、独自に整理した上で、二次的な学術的研究のために提供



## 収集データの範囲(基本的な考え方)

- **社会学に限定せず社会科学全般**(政治学、経済学、経営学、教育学、法学など)
- **個人調査だけでなく企業調査も**(後者では個人調査と異なる秘匿処理が必要に)
- **大規模調査**(調査研究機関が実施、大規模科研費によるものなど)
- **継続調査**
- **労働関係調査**(社会科学研究所の労働研究の歴史を踏まえて; 連合や産別など労働組合が実施した調査等)
- **官庁の委託調査**(統計法の対象外)など

## 寄託データの例

日本人の意識調査(NHK放送文化研究所世論調査部)  
 勤労者の仕事と暮らしについてのアンケート(連合総合生活開発研究所)  
 生命保険に関する全国実態調査(生命保険文化センター)  
 ワーキングパーソン調査(リクルートワークス研究所)  
 子ども生活実態基本調査(ベネッセコーポレーション)  
 多様な働き方に関する意識調査(内閣府国民生活局)  
 日本版General Social Surveys(大阪商業大学JGSS研究センター)  
 家族についての全国調査(日本家族社会学会全国家族調査委員会)  
 SSM調査(SSM調査研究会)  
 新規開業実態調査(日本政策金融公庫総合研究所)  
 衆議院議員総選挙の実態(明るい選挙推進協会)  
 21世紀初頭の投票行動の全国的・時系列的調査研究(JESⅢ SSJDA版),  
 2001-2005(JESⅢ研究会)  
 学生生活実態調査(全国大学生生活協同組合連合会)  
 東大社研パネル調査(東京大学社会科学研究所パネル調査プロジェクト)

SSJDAの運営状況

年 度	新規公開 データセット数 ( )内は調査数	収録調査 DB 検索 件数	公開データ リスト・ア クセス件数	利用申請 件数	利用申請 研究者数	提供データ セット数	発表論文 ・著書数	うち学位
								論文数
1998	276 (220)	-	-	11	14	22	3	1 (1)
1999	91 (60)	-	-	20	33	72	4	3 (0)
2000	42 (32)	-	-	51	74	147	3	0 (0)
2001	50 (22)	-	-	114	278	545	19	3 (0)
2002	40 (28)	4,581	-	222	489	796	37	2 (1)
2003	77 (58)	31,014	-	224	613	1,129	25	6 (1)
2004	101 (76)	40,583	-	343	802	1,735	56	12 (2)
2005	95 (65)	38,770	-	335	1,036	2,527	91	18 (1)
2006	185 (78)	43,011	-	366	944	1,641	61	19 (2)
2007	208 (94)	53,986	-	463	1,208	1,851	86	15 (0)
2008	56 (39)	43,613	5,406	481	1,434	2,165	113	30 (2)
2009	66 (62)	40,943	10,255	606	1,756	2,784	116	45 (2)
2010	61 (51)	41,425	15,425	666	1,354	1,630	135	28 (2)
2011	70 (50)	40,984	16,190	774	2,093	1,812	169	50 (3)
2012	74 (66)	40,738	16,691	855	2,567	2,352	145	74 (1)

注：収録調査DB検索件数は、2002年12月から計測

公開データ・リストへのアクセス数は2008年6月より計測

学位論文数は、SSJDA所蔵のデータを研究に利用した博士・修士・学士(卒業論文)の論文数。

学位論文数のかつこ内数は、そのうち東京大学から授与された学位論文の数。

## 2次分析研究会

- 「参加者公募型研究」: あらかじめ設定されたテーマやデータに関して参加者を募り、当センター内の研究者と共同で行うもの。
- 「課題公募型研究」: 申請者が研究テーマ・利用データ・共同研究者を設定して応募し、当センターが所有する社会調査データを用いて行う研究。

## 計量分析セミナー

- 2次分析のための統計分析手法に関する教育機会の提供
- コースの事例(「2次分析道場」)

個票データの二次分析に関する総合的な応用コースです。社研にてデータアーカイブや社会調査に携わってきたスタッフ・元スタッフが、公開データへのアクセス方法、計量分析の基礎と応用、社会科学研究的な考え方を解説します。また、実際に統計ソフトSPSSを用いてデータ分析の練習をします。SSJデータアーカイブに寄託されているリアル・データを使って、実践的な二次分析を一通り体得できることが特徴。

## SSJDA運営上の課題

- データアーカイブ利用に関する基本的なルールを利用者に徹底すること: データアーカイブは寄託者の善意で成立している組織
- 研究者が公的資金で実施したデータの寄託の促進: 多くの研究者はデータアーカイブのデータ利用には関心があるものの、自分が所蔵するデータの寄託には無関心

## 広義のデータアーカイブ間の情報共有・連携の必要性

- 運営上の課題の共有化
- データ収集の連携、データ整理やデータ提供ルールの共通化など(社会調査データフォーマットの国際標準の普及など)