

## 二次分析機能を備えたonline data archiveの利点とその要件に関する検討 — 簡易data archiveを実現するtool作成の試み —

A Study on the Merits and Requirements of Online Data Archives:  
Making of Tools to Realize Handy Data Archives on the Internet

加 藤 厚

まず、近年増加しつつある調査結果などのweb上での公開の現状の一端とその限界・制約が紹介され、続いて機関としてのdata archiveの諸機能、特に二次分析機能のinternet上での実現による諸利点及びその実現可能性の検討と作成支援toolの試作という本論文の目的が指摘された。

既存の大規模な学術的online data archiveの紹介と検討を踏まえて、本研究が実現を目指すdata archive及びその作成支援toolの要件が指摘され、それらを満たすべく試みられたJavaScriptによる支援tool (SPDA: data archiveをbrowser上に提示し二次分析を行うtool、並びにcsv2ary: 汎用data書式であるcsvをJavaScriptの2次元配列に変換するtool) の作成過程が報告された。

SPDAを用いて筆者が行ったdata archive公開の実例を提示・紹介するとともに、このような試みの意義と今後の課題について述べて本論文の結びとした。

キーワード: online data archive、二次分析、支援tool、JavaScript、情報公開

### 目 次

#### I 問題と目的

- 1 調査結果などのweb開示の現状とその制約
- 2 Data archive機能を備えたweb開示の可能性とその利点
- 3 本論文の目的

#### II 方法

#### III 結果と考察

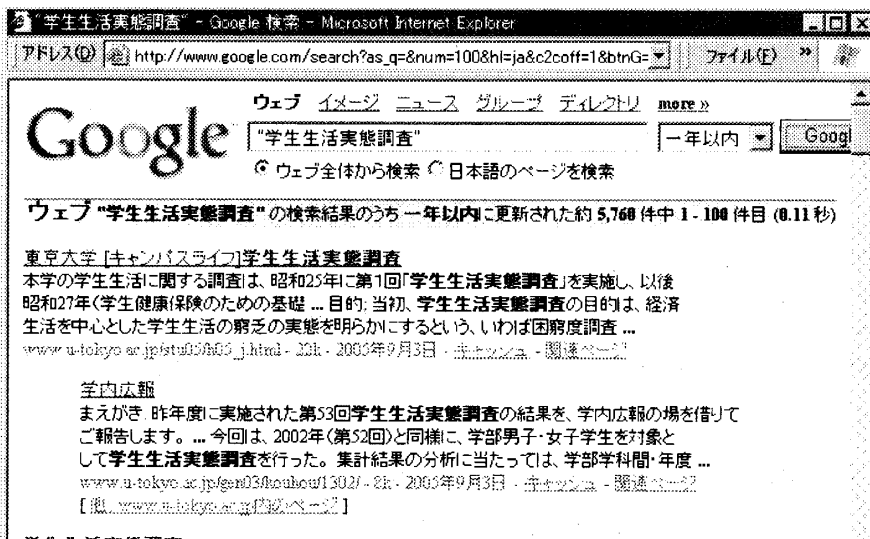
- 1 Modelとしての諸既存data archiveなどの把握と検討
- 2 目指されるべきtoolの要件
- 3 支援toolの作成とdata archive公開の過程

※本論文で解説されているspda.htm (online data archive公開支援tool) とcsv2ary.htm (data変換tool) のscriptは、資料9と資料11のURLで起動後、[表示]→[ソース]→[ファイル]→[保存]などの操作でその閲覧と保存が可能です。

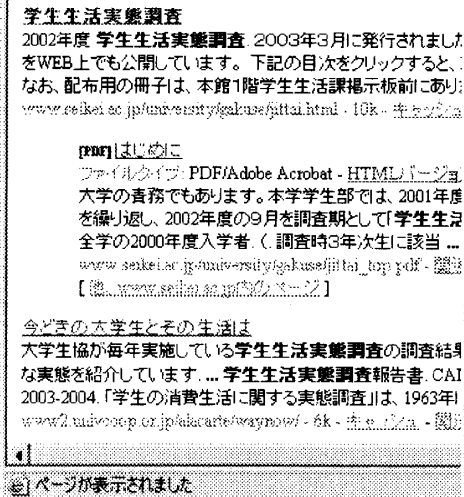
# I 問題と目的

## 1 調査結果などのweb開示の現状とその制約

近年、各種調査結果などのweb上での公開が増加している。例えば「学生生活実態調査」などの開示も、資料1に示したとおり多数存在し、比較的詳細なクロス集計結果などを含むものも多い(資料2)。しかし、このような分析及び開示は、調査実施者側の構想と便宜の範囲内でのみ行われる一方向的なものであることが多く、閲覧し利用・活用を目指す側の期待に十分応えきれない場合も少なくないことが危惧される。



資料1  
近年におけるweb上での調査結果などの開示状況の一例



学内広報 - Microsoft Internet Explorer

アドレス http://www.u-tokyo.ac.jp/gen03/kouhou/1302/11.html

1-8表 現在登録している学部・学科等(科群)に満足していますか

区分	満足している		どちらとも言えない		やや不満である		不満である		期回否	事例数	
	%	%	%	%	%	%	%	人		%	
2002年調査(52回)	(33.0)	(34.0)	(12.0)	(9.1)	(4.6)	(3.5)	(1,395)				
全体	35.4	36.0	12.3	9.9	4.3	2.0	1,501		100.0		
男子	34.3	35.9	11.9	10.8	4.8	2.3	1,142		76.1		
女子	39.7	36.5	13.6	7.2	2.8	1.1	359		23.9		
男子	前期課程	33.3	34.3	12.5	11.7	4.1	4.1	562		37.4	
男子	後期課程	35.3	37.4	11.4	9.8	5.5	0.5	583		38.6	
女子	前期課程	36.7	40.6	15.0	5.0	1.1	1.7	189		12.0	
女子	後期課程	40.8	32.4	12.3	9.5	4.5	0.6	173		11.9	
男子	文科系	35.7	35.5	10.4	10.2	6.2	2.0	501		33.4	
男子	理科系	33.2	36.2	13.1	11.2	3.7	2.5	841		42.7	
女子	文科系	36.9	35.9	14.7	8.3	3.2	0.9	217		14.5	
女子	理科系	41.5	37.3	12.0	5.6	2.1	1.4	142		9.5	

資料2  
比較的詳細な集計結果の開示の一例

調査実施者の連絡先などが明記されている場合には、閲覧者の依頼・要望により追加的な分析や開示が行われることもありうる。なぜなら、調査実施者側の着想には一定の限界があるため、立場や観点の異なる利用者からの要望などは調査実施者側にとってもしばしば有益かつ有意義だからである。しかし、調査実施者側が追加的な分析や開示を行う時間的・心理的余裕に恵まれていることは常には期待できないのが実情であろう。

## 2 Data archive機能を備えたweb開示の可能性とその利点

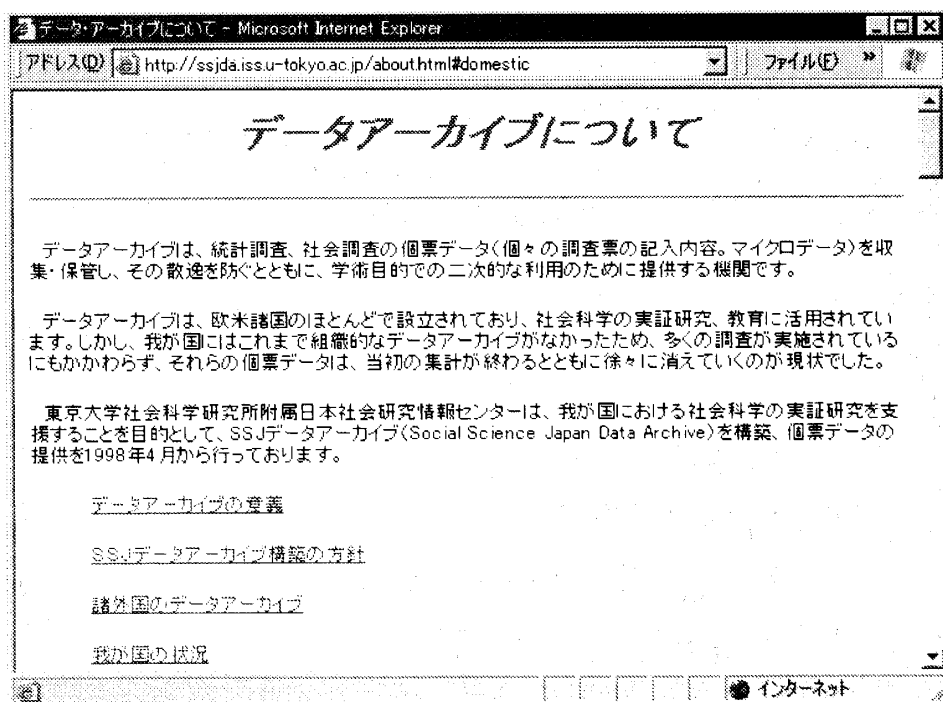
ここで、注目すべき存在にdata archive (データアーカイブ) がある。Data archiveとは、「統計調査、社会調査の個票データ (個々の調査票の記入内容。マイクロデータ) を収集・保管し、その散逸を防ぐとともに、学術目的での二次的な利用のために提供する機関」 (SSJデータアーカイブ 資料3) である。

調査結果などが開示されているweb page上に、機関としてのdata archiveが現在果たしている上記の機能、特に二次分析機能が実現できたならば、冒頭に述べた危惧はその大幅な解消が期待できる。なぜなら、World Wide Webは本来双方向的な利用に適した媒体であり、個票data分析機能がweb page上に提供されれば、開示された分析結果に飽き足らない、あるいは調査の実施/開示者側とは異なる問題意識や着想を有する閲覧者は、その興味・関心などに基づいた追加的な、あるいは新たな二次分析を自らその場で実行して結果を入手できるからである。そして、二次分析機能のonlineでの提供には、以下のような追加的利点も存在する：

- a. 貴重な原資料の有効利用 (調査対象の負担の軽減、他者による研究の継続などを含む)
- b. 多面的・包括的な分析・検討 (時系列的分析、比較研究、一次分析の検証などを含む)
- c. 実証的態度に関する広範な啓発 (仮説検証・問題解決といった教育目的の利用を含む)

## 3 本論文の目的

本論文の目的は、上記の諸利点の実現に繋がる「二次分析機能を備えたdata archive」のweb上での実現可能性を検討するとともに、それを実証する作成支援toolの試作過程を記録・報告することである。



資料3 SSJデータアーカイブによるdata archiveの定義など

## II 方法

Modelとしての諸既存data archiveなどの把握と検討、及び目指されるべき作成支援toolの要件についての考察は、web上での資料収集とその展望に基づいて行った。また、作成支援toolの試作は、JavaScript辞典（アंक、2002）などを参照しつつ筆者が行った。

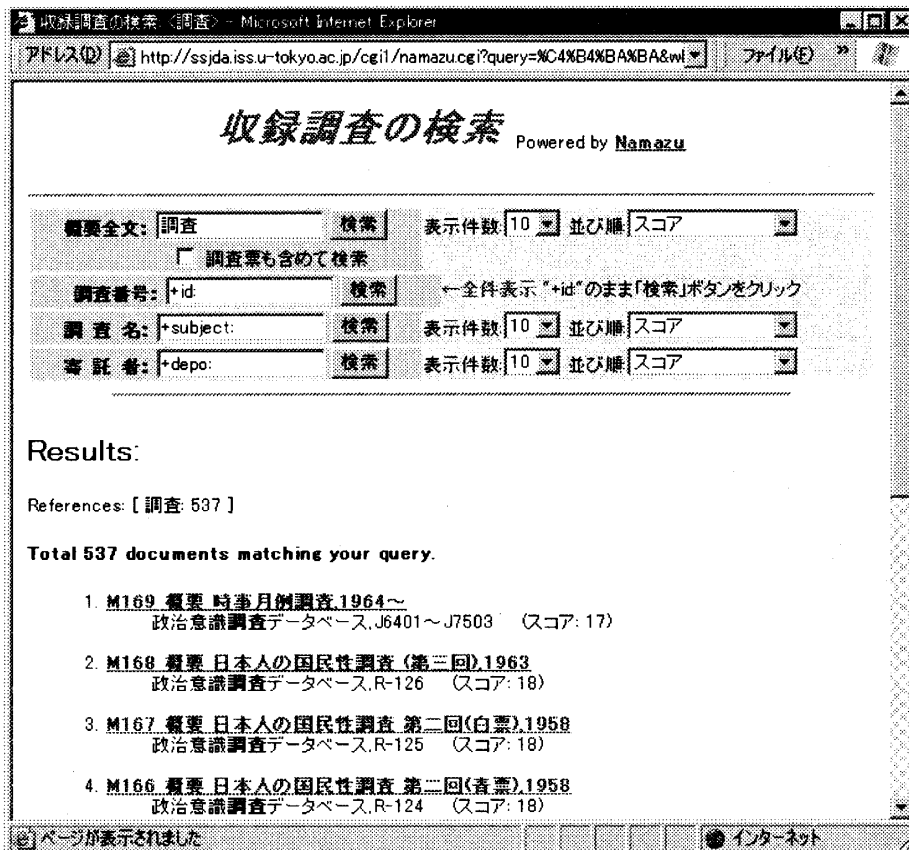
## III 結果と考察

### 1 Modelとしての諸既存data archiveなどの把握と検討

本研究でその実現が目指されるべきdata archiveの諸属性や要件を検討する手がかりとして、まずmodelとしての既存online data archiveの把握と検討を行った。検索された諸archive中、参考とすべき特長を有するものとして、以下の3件を紹介する。

#### a. S S Jデータアーカイブ

S S Jデータアーカイブ（1998年開設）は東京大学社会科学研究所附属日本社会研究情報センターが運営する国内における代表的data archiveで、2次分析が可能な調査が多数（2005年9月の時点で500件以上 資料4参照）収集・保存・公開されている点がある。



資料4 代表的data archiveであるS S Jデータアーカイブにおける検索結果の一例

他方、収集されている調査は「社会科学分野の統計調査、社会調査」にその範囲・内容が限定され、また利用者は「大学又は研究機関の研究者と教員の指導を受けた大学院生」などに限定されている。

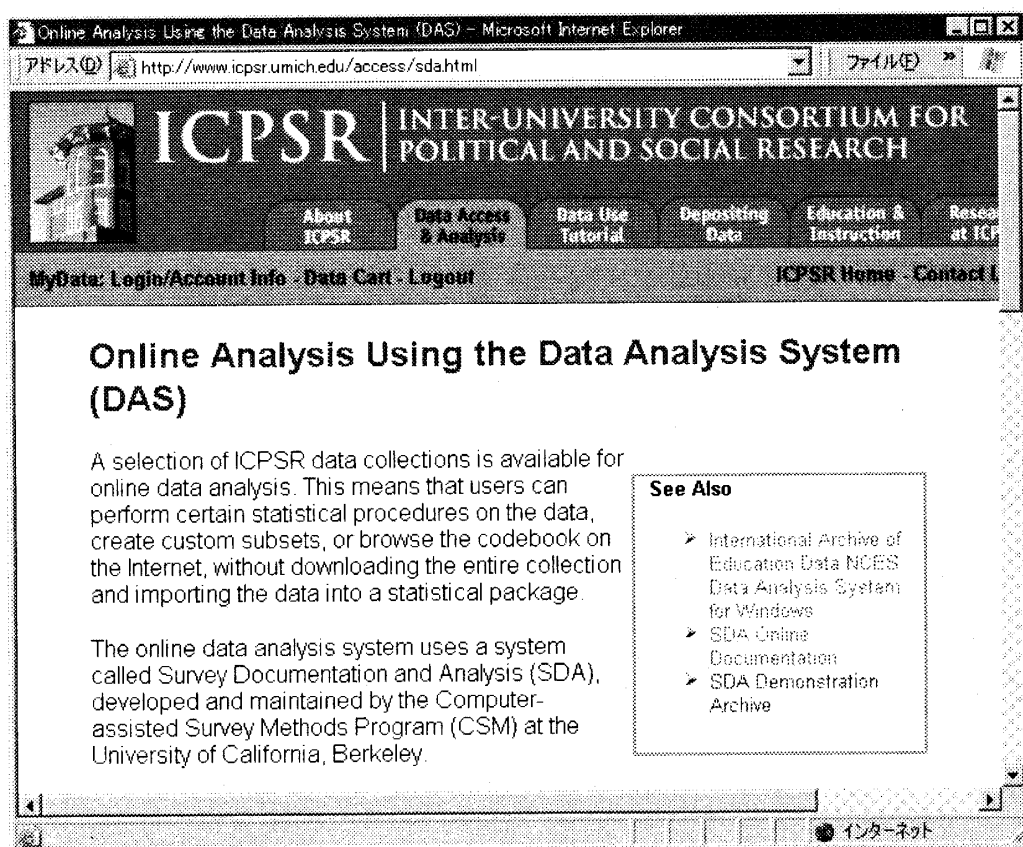
従って、例えば「学生生活実態調査」などの自発的で簡易な開示に適切な機関とは考えにくく、また全ての個人がその興味・関心の追究に自由かつ簡便に利用できるものではない。

### b. ICPSR(Inter-university Consortium for Political and Social Research)

ICPSR (1962年開設)は、ミシガン大学社会調査研究所が運営する社会科学関連調査の個票dataの世界最大級のarchiveである。

“MyData account”という「利用者登録」を行えば、公開されている調査の個票dataがdownloadでき、さらにDAS (Data Analysis System 資料5)を利用すると、収集されている調査の一部(2005年9月時点で約200 資料6)について、度数分布の集計から回帰分析に至る統計処理、読み替えや変換による新変数の作成、棒・折れ線・円などの各種graphの作図などの非常に充実した二次分析がonlineで可能な点はその特長である。

海外の研究機関であるため、身近な調査結果の自発的な開示に適切とは言えない点はSSJデータアーカイブと同様であるが、「登録」のみで広範かつ多数の調査の個票dataの入手やonline分析が可能な点は参考にすべき利点であると言えよう。



資料5 社会科学関連調査の代表的data archiveであるICPSR

## Search Results

If your search results are too broad, or if you're having trouble understanding the display of results below, you may wish to read our FAQ on [Searching the Web Site](#).

Results for: sda:SDA (199)

Document count: sda:SDA (199)

198 results found, scored using date, [sort by relevance](#) or [sort by title](#)  
sorted by relevance

1-100

### 3002 Chicago Women's Health Risk Study, 1995-1998

2000-10-18

Block, Carolyn Rebecca

[description](#) | [download](#) | [online analysis](#) | [related literature](#)

### 6399 Homicides in Chicago, 1965-1995

2005-07-06

Block, Carolyn Rebecca, Block, Richard L., Illinois Criminal Justice Information Authority

[description](#) | [download](#) | [online analysis](#) | [related literature](#)

資料6 ICPSRにおいてdata downloadやonline analysisが可能な調査の検索結果の一例

### c. SRDQ (質問紙法にもとづく社会調査データベース)

SRDQ(Social Research Database on Questionnaires) (2004年開設)は、大阪大学大学院・人間科学研究科・SRDQ事務局が運営する「質問紙法にもとづく社会調査データベース」である。2005年9月の時点で119件の調査についてその質問文並びに選択肢による検索と参照が可能であり、また2領域の計10調査についてはonlineでの二次分析も可能である(資料7)。



資料7

10調査については登録不要で二次分析が可能SRDQ

Onlineで二次分析を行う場合にも「利用者登録」などは不要であり、かつ度数分布表作成・クロス集計から因子分析・線形回帰に至る多様な分析手法が提供されている。このように、全ての利用者に対して充実した二次分析の機会を提供している点は、教育目的での利用などにおいては大きな利点である。他方、二次分析が可能な調査数については一層の充実が望まれる。

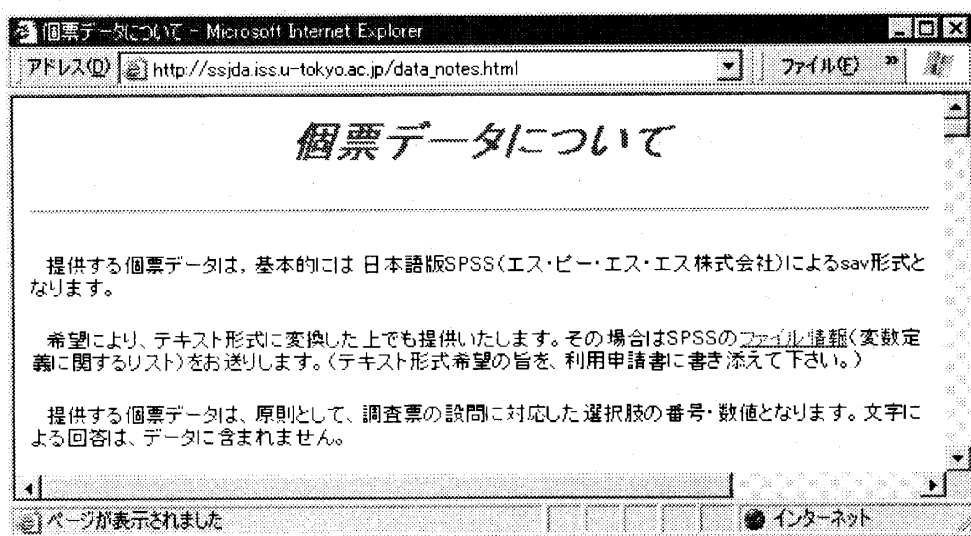
## 2 目指されるべきtoolの要件

本研究が実現を目指すのは、二次分析機能を備えた「閲覧者にとってより利用価値の高い調査結果の開示」を容易に達成しうるonline data archiveである。そこで、前節で展望した学術的で大規模な諸data archiveの特長の検討を踏まえ、本研究で試作するdata archive作成支援toolについては、下記の諸点をその要件とした。

- a. 個票dataに対する基本的な二次分析を全ての利用者が実行できること
- b. 調査に関心を持つ利用者であれば自由に使用できる程度に簡便であること
- c. 特定のOSや応用softwareを要求することなく使用できること
- d. 商用softwareの購入や専門的スキル無しにその作成と改訂が可能なこと

各要件のより詳細な内容及びその補足説明などは下記のとおりである。

a. で言及した「二次分析の実行」の必要性・意義などは「問題と目的」で述べたとおりである。なお、個票dataの範囲は「数値data」のみ、分析機能の範囲は「単純集計・クロス集計・群別集計・散布図」として開発を試みた。二次分析の対象とするdataの範囲を「数値data」のみとするのはSSJデータアーカイブと同様（資料8参照）である。また、上記の範囲の分析機能が利用できれば、2変数における質的dataと量的dataの全ての組合せの基本的分析の実行が可能である。他方、変数の変換（数値の読み替えや尺度化など）、3変数以上の分析、自由記述dataの分析などについては、今後の課題とした。



資料8 SSJデータアーカイブにおける個票data提供の形式と範囲など

また、「全ての利用者が実行できる」については、問題と目的で言及した「二次分析の諸利点」を考慮し、極力多様かつ多数の利用者における二次分析の試行を目指して要件の1つとした。

b.すなわち「利用の簡便さ」には、「二次分析を行う利用者（閲覧者）における簡便さ」と「二次分析を期待しつつ結果並びにdataを開示する提供者（調査実施／開示者）における簡便さ」の2面が存在する。前者については、mouseのみによる操作、つまりclickなどのみで二次分析の実行を可能とすることが目指された。また、後者については、data archiveの作成と公開に要する準備・技能として、web pageの公開に必要なそれら（html fileの作成とそのuploading）以外には「所定の形式でのdata file作成」しか求めないこと、及びCSV(Comma Separated Value)などの汎用data file形式からのdata file作成機能を実現することが目指された。

c.とd.すなわちOSを選ばず、応用softwareも不要で、かつ作成や改訂に高度な知識・技能を必要としない、という条件から、今回のsystemの試作はJavaScriptで行うこととした。なぜなら、まず第1に「OSを選ばない」ためには二次分析の処理はserverあるいはbrowser上で実行すること（いわゆる“web programming”）が必要であり、第2にweb programmingに使用される主な言語のうち、ActionScriptについては商用softwareが必要、またPerlやPHPについてはserverに関する、Javaについてはprogramming自体に関する比較的複雑で高度な準備や知識・技能が必要だからである。

b.の後者、すなわち「提供者における簡便さ」の追求、並びにc.d.の帰結、すなわち「JavaScriptによる作成支援toolの試作」という方針は、SRDQ（質問紙法にもとづく社会調査データベース）の現状を踏まえて指摘した「二次分析可能な調査数の一層の充実」に関する本toolの独自の貢献を期待させるものである。なぜなら、まず「二次分析可能な調査数の充実」には調査結果の自発的公開の増大が有効であり、それは公開に必要な手順・過程などが十分簡便でなければ期待しにくいからである。加えて、JavaScriptの実体はtext fileであり、そのcustomize（改造）は容易であるため、様々な「自家用変更」を加えた形でのdata archiveが公開される可能性、有志によるtool自体の改善の可能性なども期待できるからである。

### 3 支援toolの作成とdata archive公開の過程

本節では、前節で指摘した諸要件を満たすonline data archive公開支援toolの作成、及びその利用によるdata archiveの公開の過程について、下記の展開に従ってその要点を報告する。

- a. 支援tool: SPDAの概要とそれを使用したdata archiveの一例
- b. Data fileの形式並びに変換toolの作成
- c. 最適なuser interfaceの検討と実現
- d. 最適な二次分析機能の検討と実現
- e. Data archive公開の実際と今後の課題など



a. 支援tool: SPDAの概要とそれを利用したdata archiveの一例

本学の「重要問題等調査検討委員会」が実施した1実態調査の結果を開示したweb pageに呈示されたlink (左側の窓の中程) からdata archive (右上の窓) を起動し、windows XP上のInternet Explorer(以下IEと略記)で回答者の性別×出身のクロス集計 (行%付き) を実施した結果 (右下の窓) は資料9に、Mac OS X上のSafariで2種類の期待間の散布図作成 (対dataの差の検定付き) を実施した結果は資料10に示したとおりである。

The screenshot shows the SPDA (a Script for Public Data Analysis) interface. The left window displays a survey summary for 'Q2: 本学を志望するきっかけとなった [Q2A(「大学案内パンフレット」)の重視: 4]'. The right window shows the cross-tabulation results for 'Q8-性別: 1/男性・2/女性' and 'Q9-出身: 1/1市6町・2/1市6町以外の県内・3/九州他県[含沖縄]・4/その他'.

**Q2 本学を志望するきっかけとなった**  
 [Q2A(「大学案内パンフレット」)の重視: 4]

値	n	%	累積%
4	63	29.86	29.86
3	65	30.81	60.66
2	54	25.59	86.26
1	12	5.68	91.94
0	17	8.06	100.00

211 欠測 = 4 範囲(

**Item No.1 [for row/grouping/y-axis]:**  
 Q8-性別: 1/男性・2/女性

**Item No.2 [for column/stats/x-axis]:**  
 Q9-出身: 1/1市6町・2/1市6町以外の県内・3/九州他県[含沖縄]・4/その他

Frequency

Crosstabulation  Row %  
 Group means No.1(表側)×No.2(表頭)のクロス集計  
 Scatterplot  Paired t-test

Row: No.42 Q8-性別: 1/男性・2/女性  
 Clm: No.43 Q9-出身: 1/1市6町・2/1市6町以外の県内・3/九州他県[含沖縄]・4/その他

values	1	2	3	4	total
2	39 (26.897)	26 (17.931)	56 (38.621)	24 (16.552)	145 (100)
1	9 (13.235)	9 (13.235)	31 (45.588)	19 (27.941)	68 (100)
total	48 (22.535)	35 (16.432)	87 (40.845)	43 (20.188)	213 (100)

null=2 coefficient of association(Cr)=0.19355  
 chi-square=7.98 p<.05 Yates' chi-square=6.45 p<.1  
 minimum fe=11.174 fe<=5 cell(s):0%

資料9 IEで表示されたweb pageのlink (背景=左側windowの中程) から data archiveを呼び出して行った2変数間のクロス集計の一例

本tool: SPDA (a Script for Public Data Analysis) は上に示した両browserに加えて、例えば Netscape、FirefoxといったJavaScriptに対応する全てのinternet browser上で作動し、user interface的にはpointerによる選択とclickにより全ての操作が可能である。また[help] (右上の窓の下部)、各buttonの説明などはpop-up形式で表示される (資料9の右上の窓)。

1. 新入生対象の「本学への志望動機等に関する基本調査」の調査票の投票期に提示されています。  
また、buttonをクリックすると、より詳細な集計/図示が

a. 調査実施：4月22日と27日に11ゼミで、5月6日と11日  
b. 有効回収：計215名(11ゼミ：181名+2ゼミ：34名)  
c. (内訳+調査票check)

2. 下記以外に必要な集計・図示などがありましたら、request

d. 情報源の重複度 (Q3) :     
e. 本学志望理由 (Q3) :     
f. 期待度 (Q4) :   →追加(5/31)  
\*追加分析 :  →追加(6/3)

3. 本調査の数値dataを archive 化しました。link+追加  
群間での変数や平均値の比較、相関係数の算出などの二次分析も

May 15, 2005

Q2 本学を志望するきっかけとなった情報源

{Q2A(「大学案内パンフレット」)の重視: 4/大いに3/かなり}

頻	n	%	累積%	0	10
4	63	29.86	29.86	*****	
3	65	30.81	60.66	*****	
2	54	25.59	86.26	*****	
1	32	15.69	91.94	***	
0	17	8.06	100.00	*****	
-----					
211 欠測 = 4 範囲(0~9999)外 =					

Item No.1 [for row/grouping/y-axis]:  
Q4-3A-学習・研究環境の充実を期待: 全10点から配分

Frequency

Item No.2 [for column/status/x-axis]:  
Q4-3C-部・サークル活動の充実を期待: 全10点から配分

Crosstabulation  Row %  
Group means  
Scatterplot  Paired t-test

名称未設定

No.33 Q4-3A-学習・研究環境の充実を期待: 全10点から配分 mean=3.443

```

[10] 1
      2
      3
      4
      5
      6
      7
      8
      9
      10
[0] 2
      1
      2
      3
      4
      5
      6
      7
      8
      9
      10
[0] ----- [7] -----

```

No.35 Q4-3C-部・サークル活動の充実を期待: 全10点から配分 mean=2.148

valid n=210 null=5 r(208)=-0.4436 p<.01  
t for paired data(209)=6.134 p<.01

資料10 Safariで表示されたweb pageのlink (背景=左側windowの中程) から data archiveを呼び出して行った散布図作成 (対dataの差の検定付き) の一例

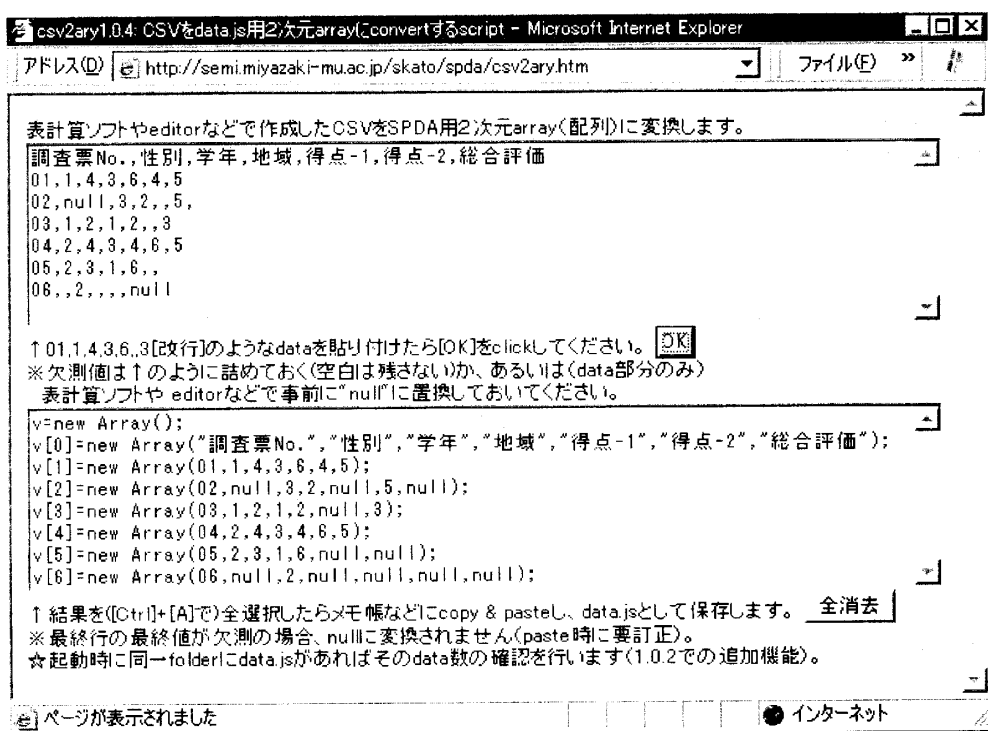
単純集計、クロス集計、群別集計、散布図といくつかの追加的統計機能の装備、pointerとclickのみでの操作、JavaScriptによる作成とその結果としての殆どのinternet browserへの対応、及び容易な改造可能性などにより、「2. 目指されるべきtoolの要件」で示した各要件はほぼ充足されていると言えよう。

b. Data fileの形式並びにdata変換toolの作成

変数間を半角カンマで区切るCSV (Comma Separated Value) 形式は最も一般的なdata fileの形式の1つである。この形式で作成したfileをJavaScriptで処理するためには、2次元配列 (JavaScriptではArray) に格納したdataを外部JavaScript fileとして読み込むのが最も合理的であろう。また、data archiveの自発的な公開が進展するためには、「2. 目指されるべきtoolの要件」で指摘した

ように、必要な手順・過程などの簡便さが重要である。そこで、既存dataを2次元配列の外部JavaScript fileに変換するtool(csv2ary.htm)をJavaScriptで作成した。

資料11に示したように、欠測値を含むCSV fileを上部のtext areaに貼り付けて[OK]をクリックすると、まずnewステートメントで新配列vを宣言し、続いて第1行のdata(変数名)は""で括り、第2行以下のdata(変数値)は欠測をnullに置き換えて、vの要素内に順次追加してゆくscriptが下のtext areaに表示される。



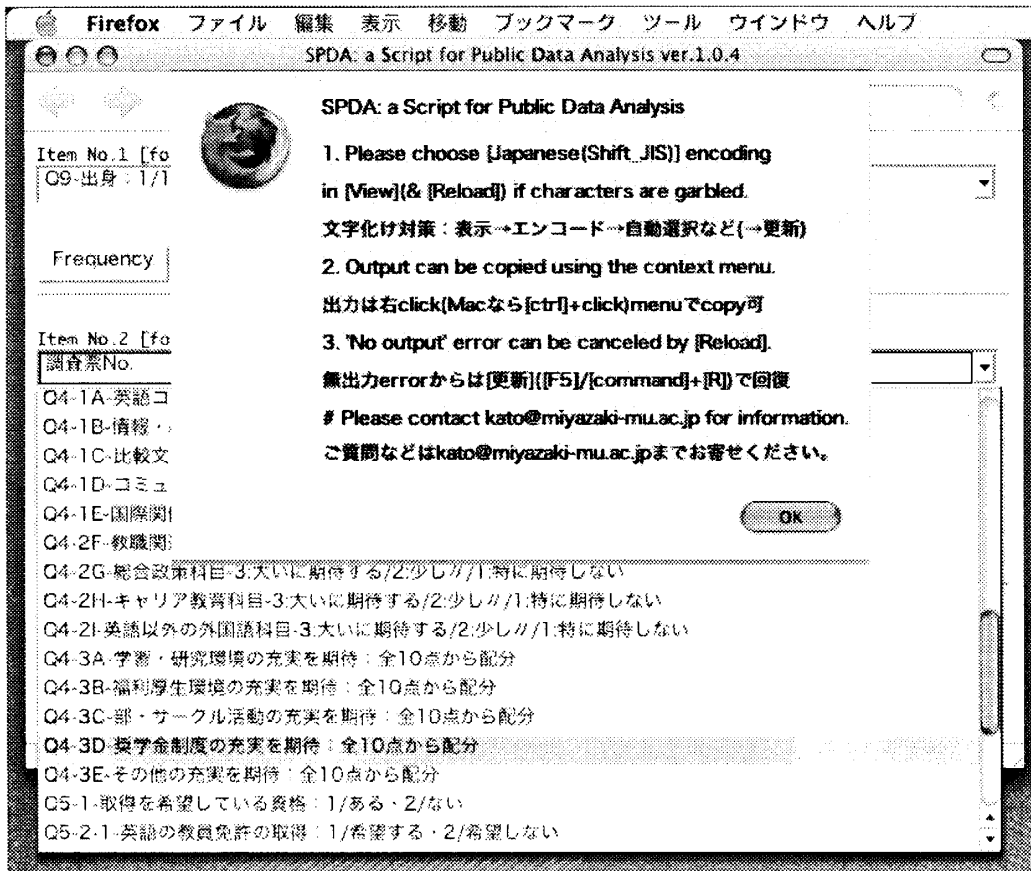
### 資料11

一般的なdata形式であるCSV fileを貼り付けてSPDA用の2次元配列を得るためのtool(csv2ary)の実行例

### c. 最適なuser interfaceの検討と実現

現在、標準的なuser interfaceはGUI(Graphical User Interface)であり、SPDAにおいても二次分析の対象とする変数や実施する分析の選択は、pointerのみで行えることが望ましい。幸い、internet browserの標準的なGUIであるプルダウンメニュー、ボタン、チェックボックスなどは全てJavaScriptによって利用可能である。そこで、選択肢の数が百以上になりうる変数の選択についてはプルダウンメニューを、数種類に限定される分析の選択についてはボタンを、そしてon/off択一の追加的統計機能の実行の有無についてはチェックボックスをinterfaceとすることとした。また、資料9に示したように、ボタンやチェックボックスにpointerを重ねると、各々の機能に関する日本語の補足説明が、そして資料12に示したように、[help]にpointerを重ねると、SPDAに関する一般的な補足説明がpop upするようにした。

なお、一般的なinternet browserにおいては、GUIに加えて各種のkeyによる操作(例:[Tab]や[↓]による操作対象の変更、[space bar]によるチェックのon/off、[Enter]による確定・実行)も可能であるが、その有効範囲は各browserに依存する。



資料12

変数選択にプルダウンメニューを、helpなどにpop-upを用いたSPDAのuser interface

※Mac OS X上のFirefoxでの実行例

※プルダウンメニューとhelpの同時表示は実際には不可能

d. 最適な二次分析機能の検討と実現

結果と考察の「1 Modelとしての諸既存data archiveなどの把握と検討」で展望したdata archive中、onlineでの二次分析が可能なのはICPSRとSRDQである（2005年9月現在、SSJDAは「リモート集計システム」を学内限定で試験運用中）。

ICPSRのDASで利用可能な二次分析機能は、度数分布集計、クロス集計、平均値の比較、相関行列、回帰分析などであり、加えて個別ケースの値の一覧や変数の変換(Recode & compute)、そして度数分布集計とクロス集計については棒・折れ線・円グラフ(bar, line, and pie charts)の作図も可能である。なお、DASはUniversity of California, BerkeleyのCSM(Computer-assisted Survey Methods Program)が提供するSDA(Survey Documentation and Analysis)を使用している。

また、SRDQで利用可能な二次分析機能は、度数分布表、クロス集計、一元配置の分散分析、相関分析、偏相関分析、因子分析、線形回帰などである。加えて、条件・関数などを用いた「ケースの選択」も可能である。なお、SRDQはSPSS Japanの“SPSS Webapp Framework”を使用している。

本研究で作成するSPDAにおいては、「2. 目指されるべきtoolの要件」で述べたとおり、「2変数における質的dataと量的dataの全ての組合せの基本的分析の実行が可能」な「単純集計・クロス集計・群別集計・散布図」の範囲での開発を試みた。加えて、単純集計については累積%・

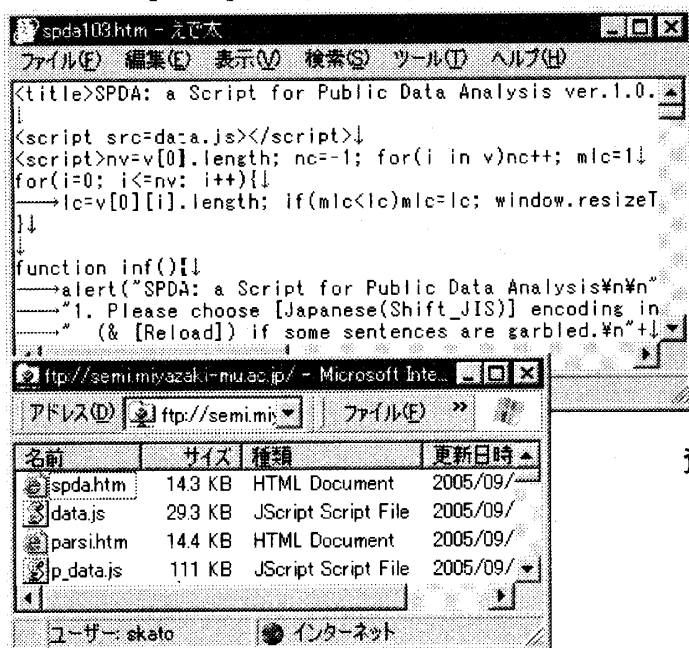
標準偏差など、並びにヒストグラム／横棒グラフの提示、クロス集計についてはクラメールの連関係数Cr、カイ二乗値とその有意性の検定、並びに最小期待度数(fe)など、群別集計については相関比、並びに一元配置の分散分析とその有意性の検定、そして散布図については相関係数とその有意性の検定、並びにdataの差の相関 t 検定などの実現を目指した。

これらの機能は、筆者の作成したQPD(Questionnaire Processor for DOS 加藤、2004)ではDOS汎用BASICであるTBC(TOKIWA BASIC Compiler 岡田、1998)によって既に実現されている。そこで、同様のalgorithmがJavaScriptによって再現された。

### e. Data archive公開の実際と今後の課題など

このようにして作成されたSPDAは約14KBのhtml file(spda.htm)であり、server folderにuploadして調査結果の概要を開示したweb pageなどからlinkしておけば、そのlinkのclickによって同一folder内に存在するdata file: data.js (JavaScriptの外部ファイル)を読み込んで起動する。資料9・10に示したのがその状態である。

SPDA内にはdata fileの名称がdata.jsと記述されているため、scriptを一切変更することなく複数のdata fileを開示するには、それぞれ内容の異なるdata.jsをSPDAとともに異なるfolderにuploadする必要がある。ただし、data fileの名称は資料13に示したとおりspda.htmの冒頭部分(第3行)に1箇所記述されているのみなので、その部分を書き換え、かつSPDA自体の名称も変更すれば、1つのfolderでの複数のarchiveの公開は容易に可能である。例えば、資料14に示したように、筆者は2005年9月現在、同一のserver folderに2つのdata archiveを保存・提示している。この場合、一方のdata file (p\_data.js)の開示用SPDAは別名(名称は任意、このdataの場合は“parsih.htm”)とし、その第3行では読み込まれるべきdata fileとしてp\_data.jsが指定されている(第3行の具体的記述は<script src=p\_data.js></script>)。



資料13 SPDAのscript (spda.htm)の冒頭部分

※Data file名(data.js)は、外部JavaScript fileとして第3行に指定されている。

資料14 Scriptのわずかな書き換えで同一folderでの複数data archiveの開示も可能

今回試作された簡易data archive公開支援tool: SPDAの今後の課題としては、下記の3点を指摘したい。

1. Online data archive公開の試行の継続による経験と成果の蓄積
2. 必要かつ有効な範囲内の「二次分析機能」などの追加
3. Scriptの一層の洗練

「日本では、特定分野に関するデータアーカイブは存在したが、社会科学分野のデータを包括的に収集、整理し、2次分析に提供するデータアーカイブは、最近まで設立されてこなかった」との指摘(佐藤、2003)が示すように、online data archiveの試みは、特に国内においてはまだその黎明段階にある。そして、このような試みの進展・展開には、研究・調査の方法の共有とその結果としての改善、既存の結果の第三者による検証とその結果としての再現性の確認といった効果が期待できる。つまり、online data archiveは、以下の引用の中で筆者がかつて指摘した「知識の民主化」をさらに一歩進める上で有効な1手段なのである。

「大学生あるいは大学院生が研究法を学び、自分なりのデザインで実験・調査を行い、その結果を処理してある知見を得、発表することの意義は何でしょう？ それは一言でいえば『知識の民主化』だとぼくは考えます。(略)研究法を学び、実践することによって、ぼくたちは『権威への服従』に依ってではなく、理性的な合意の下に知識を得る1つの方法・1つの態度を身につけるのです。そしてこれは研究者にのみ必要なものではありません。」(加藤、1988)

これらの効果や利点、意義などの存在は、課題1.「公開の試行の継続」の必要性を支持するものといえる。

課題2.「機能の追加」に関しては、結果と考察の「2. 目指されるべきtoolの要件」で述べたとおり、変数の変換(数値の読み替えや尺度化など)、3変数以上の分析、自由記述dataの分析などが今後の課題として想定される。特に、基本的な「data変換機能」(読み替え、足し上げなど)は、「分析対象の限定機能」とともに、より柔軟で多面的な二次分析を可能とするためにもその必要性は大きい。

課題3.については、筆者の知識・技能の不足により、spda.htmとcsv2ary.htmの両toolの動作は、help(資料12参照)の3.などでも言及しているとおり完璧とはいいがたいのが現状である。全ての利用者がそのsource(programの実体)を閲覧・改良できるJavaScriptの特長を活かし、利用者・読者諸賢からご教示・ご助言などを頂ければ幸いである。

## 文献

(株)アंक 2002 JavaScript辞典 翔泳社

<http://www.seshop.com/detail.asp?pid=3126>

岡田 元次 1998 TBC(TOKIWA BASIC Compiler)

<http://www.vector.co.jp/soft/dos/prog/se031410.html>

<http://www.ehime-u.ac.jp/~achem/solid/tbc/tbc.html>

加藤 厚 2004 教育実践を支援する基本的統計処理toolの必要性並びに現状とその1例としてのQPDの過去・現在・未来

宮崎公立大学人文学部紀要 第12巻 第1号 pp. 45-62.

<http://semi.miyazaki-mu.ac.jp/skato/qpd-a.doc>

加藤 厚 1988 パッケージ野原でかくれんぼ 教育情報処理演習報告書 第5巻

筑波大学教育研究科障害児教育専攻 p. 58.

佐藤 博樹 2003 計量的な社会科学研究とデータアーカイブの機能

大学時報 288 社団法人日本私立大学連盟

<http://ssjda.iss.u-tokyo.ac.jp/news200301.html>