

## 第4回 SPARC Japan セミナー2018

「人文社会系分野におけるオープンサイエンス ～その課題解決に向けて～」

# 人文学・社会科学データインフラストラクチャー 構築推進事業の取組について

前田 幸男

(日本学術振興会 人文学・社会科学データインフラストラクチャー  
構築推進センター / 東京大学大学院情報学環)

### 講演要旨



2018年4月より日本学術振興会は人文学・社会科学データの保存と共有を推進するための活動を行ってきた。本事業の目的は今まで蓄積されてきた社会調査・社会科学データを保存し、さらなる分析のために共有する基盤を整備することにある。また、データの利活用を促進するための環境整備もその活動の一環である。現段階までの具体的活動は、既存のデータ保存・共有活動の強化と、データの英語化・国際化の支援である。さらなる環境整備は今後の課題である。

また、それらのデータの有効な利用を促すためにも、データ利用者が効果的に研究目的に沿ったデータを発見できるようにする検索システムの構築が重要な課題である。海外の社会科学データ・アーカイブを参考にしながら準備を進めているデータカタログの計画についても紹介する。



### 前田 幸男

日本学術振興会研究員／東京大学大学院情報学環教授。東京大学法学部卒業(1993)、同法学政治学研究科修士課程修了(1995)。米国ミシガン大学政治学部博士課程修了(Ph.D. in Political Science, 2001)。東京都立大学法学部／首都大学東京を経て、2006年に東京大学社会科学研究所・大学院情報学環着任。2016年より同教授。主な研究テーマは世論調査と投票行動。大学院生の頃から社会調査データの保存と共有に関連した研究にも取り組んできた。

私は東京大学大学院の情報学環に所属していますが、日本学術振興会でも研究員をしているため、本日は日本学術振興会の研究員という立場で報告をさせていただきます。私の本職は政治学者で、普段は世論調査のデータ取得・分析をしています。今日は私の考えというより、学術振興会でどういうことをやっているかということをご報告します。

### 人文学・社会科学分野の取り巻く環境

人文学・社会科学は、多様な文化や価値観に対する認識を深め、社会的な合意形成や社会的対立の解決方

法を探求する際に非常に大きな役割を持ちます。ここで、社会調査データや数量的統計データを保存することがなぜ重要なのかということをもう一度確認したいと思います。一般的に古文書や書簡、公文書が歴史研究において非常に重要であって、そのようなものをきちんと保存する活動がずっと行われてきたことは周知のことですが、社会調査データや数量的統計データも、文字ではありませんが、数字としてその時代の社会の状態を記録した極めて重要な資料です。ですから、文字であろうが地図であろうが書簡であろうが統計データであろうが、非常に重要なデータ・資料であるとい

う意味では共通しているのです。

われわれは今、社会的あるいは経済的な課題に直面しており、少子化や晩婚化、社会・経済的不平等の拡大がよくいわれますが、これらの課題を、学術目的だけでなく政策目的でも検討しようとした場合、今の段階で社会・経済的不平等が拡大していることの原因をきちんと究明するためには、社会・経済的不平等が拡大する以前のデータが必要です。われわれが直面する重要な社会的課題について理解を深めることは、きちんとデータが残っていなければ不可能なのです。ですから、単純に現在のために研究上の資料をきちんと保存しましょうという話ではなく、今の研究活動から生み出されているデータは、少なくとも人文・社会系のデータに関しては、30年後、50年後の研究者が時代を振り返って研究するときのための大事な資料になることはほぼ間違いないので、社会調査データ・数量的統計データを保存し、公開していくことは極めて重要な活動であるということなのです。

その上で、さらに最近では、オープンサイエンスの動向もあり、データを保存し、みんなでシェアすることの重要性も強調されるようになりましたし、政策立案においても、エビデンスに基づいて政策立案するためには客観的・数量的なデータが必要だという考え方が共有されるようになってきました。そのような活動の一環として、日本学術振興会でも、データを保存し、共有し、皆さんで利用することを進めていくという趣旨で活動しています。

諸外国において社会科学データを共有する取り組みが組織的に行われるようになったのは基本的には1960年代からで、海外の大きな、有名なデータリポジトリ、データアーカイブは、1960年代に活動を始めているものが多いです。イギリス、ドイツ、アメリカ、みんなそのころに始まっています。時代によって違いや変遷はあるのですが、そのような意味で長い活動の歴史があります。

それに関して日本では、1980年代ぐらいからさまざまな構想や取り組みはあったのですが、最低限の活

動が行われるようになったのは1990年代からではないかと思います。それ以前はさまざまな重要な標本調査が行われてはいましたが、それらの取り組みは基本的に研究者のグループや各研究室で保存・共有するという努力が行われ、組織立った、体系立った政策的支援が行われてきたとは言い難い状態です。特に、現在も古いデータはまだ死蔵されているものがあり、そういったデータはそろそろ今、電子化して保存するという手続きを踏まなければ、下手をすると未来永劫失われてしまう状態にあります。実際問題として失われたものも既にあるかと思います。

日本は人口規模も大きい国なので、日本語だけで学術活動ができるという意味では、これは日本という国が持っている強みでもあるのですが、なかなかデータや資料の英語化が進まず、やはり国際発信が少し弱いところがあります。ヨーロッパの人口規模が小さい国では、英語が母語でなくてもほとんどのものが英語になっていることも珍しくないで、日本は国際発信に若干後れを取っています。社会科学に関してもデータを英語化して共有することで、海外の研究者が日本のデータを分析しやすくなると同時に、日本の研究者が海外で研究報告をすることもやりやすくなると思われ

ます。そのような努力は他のアジア諸国でも進んでおり、このままデータを共有する努力や国際化する努力をしなければ、ますます社会科学の振興は遅れてしまうのではないかという危機感から、日本学術振興会は、平成30年度より、「人文学・社会科学データインフラストラクチャー構築プログラム」を新設し、社会科学データの共有・利活用を促進する基盤を構築する活動をしています。私はそのお手伝いをしています。

## **人文学・社会科学データインフラストラクチャー構築プログラムでの取り組み**

図1は、本事業のスローガンのようなスライドです。日本の社会科学全体でこのデータ共有に取り組み、学術振興会を中核に、重要かつ研究者の利用に供するこ

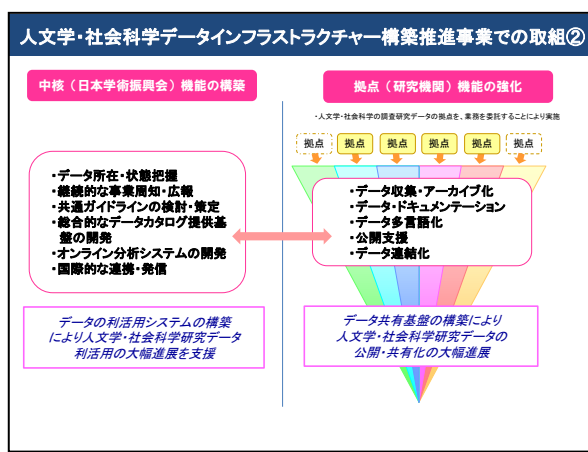
とが有用である学術調査データの保存・管理の取り組みを行う幾つかの拠点を形成すると同時に、それらの拠点をつなぐような活動を学術振興会で推進し、分野や国を超えて社会調査、人文学・社会科学のデータを共有する仕組みをつくるのが本事業の目的です。

ここからは個人的見解ですが、このような仕組みをつくる時、一つの組織に全部を集めるというやり方もあります。ヨーロッパの人口規模が小さい国では、そのようなやり方をしていることが多い印象がありますが、私が見る限り、人口規模が大きい国では、むしろ複数の組織のアライアンスを組んで活動することが多いように思います。本事業の活動はどちらかというところアライアンス型ということになります。

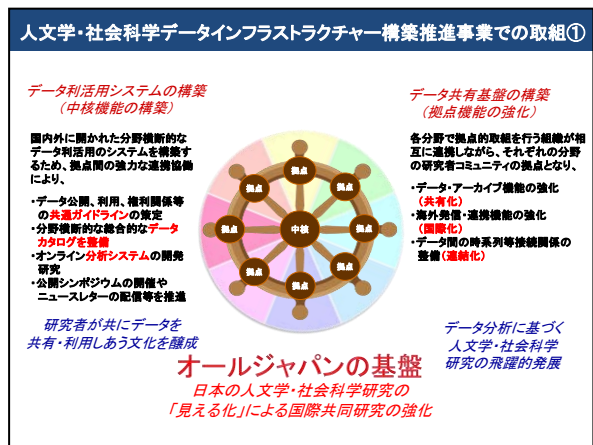
本事業の具体的な活動について紹介します。活動には二つの種類があると考えています。図2の左側が学術振興会あるいは学術振興会と一緒に活動を行う機関の守備範囲で、右側が、実際により具体的なデータに関する活動を行う機関の活動内容です。右側では、拠点機関である大学が、既存のデータを集めて、それを保存する手続きをします。また、データだけでは意味を成さないでそのデータを理解するための十分なメタデータを準備する、データを多言語化するということを行います。

その一方、日本学術振興会では、もちろんそのようなデータを発掘する作業もするのですが、むしろ共通のガイドラインを定めたり、政策的な努力をしたりします。それと同時に、各拠点機関が集めたデータのカタログを連結し、データをできるだけ可視化して、皆さんがどこにどのようなデータがあるか分かるようにする仕組みをつくることを重要なミッションとしています。

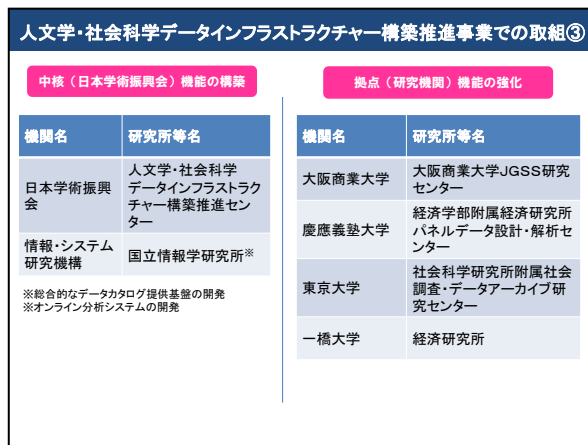
図3は具体的な陣容の図です。中核を担うのが日本学術振興会の人文学・社会科学データインフラストラクチャー構築推進センターで、現在はセンター長、副センター長の他、研究員が2人います。研究員のうち、1人が私で、私はどちらかというと社会学と政治学のデータをカバーするのに対し、もう1人は経済学のデータをカバーするといったように、守備分担をしています。データの保存・提供の具体的な活動は、右側に示してある四つの大学の附置研究所なり研究センターが担います。



(図2)



(図1)



(図3)

## 人文学・社会科学系学術的調査データの現状 に関する調査

次に、私たちが今年の5月から6月にかけて行った、人文学・社会科学系学術的調査データの現状に関する調査の内容を簡単にご紹介します（図4）。大学のリポジトリから参加していらっしゃる方は、このあたりの話が一番面白いのではないかと思います。これは現在、学術調査データがどのような形で共有・保管されているのかについて具体的に情報収集し、今後の事業の基礎資料とするために行ったものです。

調査対象者は国内の大学などに所属する人文学・社会学分野の研究者で、過去20年間に科研費を採択されたことがある方に大学を通じて調査への協力をお願いしています。大学を通じて、きちんとしたリストから調査へ協力をお願いし、それに対してインターネットを通じて回答いただくという形で調査しました。回答にご協力いただいたのは、3,956人の方々です。

内容を紹介します（図5）。まず調査票の保管状況です。ここで「N」とあるのは、人数ではなく、データセットの数です。これは、1,517人が「自分が保管している」と答えたのではなく、存在する2,596のデータセットのうち、1,517データセットについては研究者本人が保管していると読みます。1人でたくさんのデータを集めている人がいるので、個人単位の集計ではなく、データファイル単位の集計だと考えてください。

そうすると、調査票の58.44%は研究者個人が保管

人文学・社会科学系学術的調査データの現状に関する調査①	
○ 目的: 人文学・社会科学の分野における学術的調査データの現状や保管状況等について基本的な情報を収集し、今後の事業展開のための基礎資料とする。	
○ 調査対象者: 国内の大学等(約1900機関)に所属する人文学・社会学分野もしくは、当該分野に関わる学際的な領域において学術的調査を実施し、調査データを収集・分析して学術研究を推進した研究者。 (特に過去約20年間に科研費に採択されたことのある研究代表者)	
○ 調査方法: インターネットを利用したオンライン調査形式	
○ 調査時期: 平成30年5月7日～平成30年6月20日	
○ 回答者数: 3,956	

(図4)

している状態になります。22.34%は廃棄されています。電子化されて廃棄されているものはいいのですが、素データを電子ファイル化しているかを見ると、2,596のうち1,674に関しては電子ファイル化されているけれど、912については電子ファイルの状態ではなく、これは恐らく紙の原票の状態だと思われます。

より重要なのは図6です。データ公開の状況を見ると、2,596あるデータセットのうち、13.7%は公開されており、6.0%は公開準備中なので、20%程度のデータは公開されている、あるいは公開の予定であるということですが、逆に80%ぐらいのデータは公開していない、共有していない状況です。

公開されていない2,073データセットの公開の可能性を見ると、「可能」が79、「公開に向けたサポートがあれば可能」と研究代表者が言っているものが512で、一方、「不可能」も920、44.4%あります。

人文学・社会科学系学術的調査データの現状に関する調査②				
調査票保管状態		素データを電子ファイル化しているか		
保管状態	N	%	電子ファイル化	
自身が保管している	1517	58.44	されている	1674
他者もしくは機関が保管している	343	13.21	されていない	912
廃棄した	580	22.34	不明	10
その他	66	2.54	計	2596
不明	90	3.47		
計	2596	100.00		

(図5)

人文学・社会科学系学術的調査データの現状に関する調査③					
データ公開の状況			データ公開していない場合、公開の可能性		
データ公開の状況	N	%	データを公開する可能性	N	%
公開中	355	13.7	可能	79	3.8
公開準備中	156	6.0	公開に向けたサポートがあれば可能	512	24.7
公開していない	2073	79.9	不可能	920	44.4
不明	12	0.5	不明	562	27.1
計	2596	100.0	計	2073	100.0

(図6)



「公開に向けたサポートがあれば可能」というのは、具体的には、恐らく、データのクリーニング・ドキュメンテーションや、提供の窓口を他が代行してくれるのであればできるという意味であり、先述の、人文学・社会科学データインフラストラクチャー構築プログラムでの拠点機関の活動にそのような支援があれば、これは解決できるのではないかと思います。

それに対して「不可能」に関しては、その理由を事細かに聞かなければ分からず、もちろん単にやりたくないという人もいるかもしれませんが、私が推測するに、「分からない」という理由が多いと思うのです。例えば、調査対象者の同意をきちんと取っていないかった、あるいは、ものすごく人口の小さい、例えば東京都の八丈島で調査が行われていると、秘匿処理の問題があって調査データを出せないのです。そうすると、どの程度まで秘匿すれば出せる状態になるのか、個別の研究者では分からないということが十分考えられます。そうすると、例えば事前に「このような挨拶状を作ればデータは問題なく共有できる」「この人口をここまで大きい単位に丸めたら基本的に公開しても個人が特定されることはない」というようなガイドラインが明確になれば、データを出すことができるのではないかと思います。その観点から言うと、こちらの不可能を可能にするためには、本事業で学術振興会が担う中核機能の構築で、ガイドラインや方針を決めれば、全部とは言いませんが、好転する部分はあるのではないかと考えています。

### **拠点機関の活動**

次に、各拠点機関の活動の取組について紹介します(図7)。求められる取組の一つ目は、データアーカイブ機能の強化(共有化)で、データをたくさん集めてドキュメンテーションを作り、多くの方に使っていただきます。二つ目は、海外発信・連携機能の強化(国際化)で、データを多言語化し、できるだけ国際的に日本のデータを使っていただけるようにします。三つ目は、データ間の時系列等接続関係の整備

(連結化)で、データの利便性を高めます。これは少し分かりにくいかもしれませんが、統計データは取られた時代によって基準が変わることがあります。分かりやすい例を挙げると、自治体を単位にして作られたデータでは、選挙区制度改革の以前と以後で選挙区が違います。そうすると、選挙の結果について時系列で長期比較するようなデータファイルを作ろうとすると、区域を分けて接合し直さなければいけません。選挙に限らず、社会経済統計データでも、平成の大合併以降、自治体の区域は随分変わっているのです。そのようなことを個別の研究者が逐一、ばらばらでやっているという極めて非効率的なのです。そのように、国民経済計算の基準が変わったときの接続という問題もありますが、さまざまところで組織的に努力すればみんながハッピーになって使えるけれど、個人がばらばらにやっているというとても非効率な作業はたくさんあります。そのような活動に協力・支援する必要があると考えています。

本事業のデータの範囲は、社会科学分野の研究のために取得・生成されたものであって、主に、人々の意識や態度、状態、行動などを調査、記録した電磁的記録とし、個票データの他、集計データ、時系列データおよび質的データ等も含まれます。結果として、基本的に数値のデータが多かったですが、さまざまな集計データやアグリゲートデータ、質的データも念頭に置いて活動を開始しています。

取組は、初年度については、まず手の付けやす

拠点機関の活動①	
1. 求められる取組	
a. データアーカイブ機能の強化(共有化)	
b. 海外発信・連携機能の強化(国際化)	
c. データ間の時系列等接続関係の整備(連結化)	
2. データの範囲	
社会科学分野の研究のために取得・生成されたものであって、主に、人々の意識や態度、状態、行動などを調査、記録した電磁的記録とし、個票データの他、集計データ、時系列データ及び質的データ等も含まれます。	

(図7)

いところからということで、各研究所やセンターが、自機関が所有するデータについて優先的に作業することが可能ということで、拠点の活動を始めていただいています。具体的な取り組み例は、「共有化」については、データを長期的に保存・共有するためのシステム構築・環境整備、データの補定やウェイトの作成、個体識別の可能性が高いデータを分析する環境の整備（秘匿処理など）、データカタログの整備、データと研究成果・論文等との関連づけ、DOIの付与があります。「データの補定」というのは、分かりやすい例では、世論調査データなどで、所得を答えてくれない方がいらっしやるのです。それはいいのですが、そこを欠損のままにしまうと、国際比較調査などでは、「そのデータは不完全だからきちんと完全なものに出してください」と言われることがあります。そのときに、統計学の世界で、データインピュテーションという、統計学的にどういう形でその欠損値を埋めるかという議論があるのですが、そのような加工情報をきちんと付けた上でデータをリリースすることが分野によって必須になっていて、データの補定ではそのような作業をします。

「国際化」については、国際比較調査データの作成およびメタデータの作成、データおよびメタデータの英語化、海外機関との連携した活動という取り組みがあります。データおよびメタデータの英語化は、外国の研究者も見られるようにするためですが、英語化によって、外国のジャーナルに投稿する分析にそのまま比較的利用しやすい形にもなる、ということをしています。

「連結化」については、政府統計データの時系列接続、標本調査データの累積データファイルの作成が挙げられます。累積データファイルを作るのは非常に手間と暇が掛かりますが、一回できれば誰にでも簡単に、すぐ時系列の表などが作れるので大変便利なのです。そのような活動を組織立ってやる方が、研究、アウトプットを増やすという観点で有益なので、そのような活動も拠点機関にお願いしています。

2年目以降については、他機関が保有する既存データについても収集を開始して、以上のような取り組みを実施してもらいます。

拠点機関に取り組みをお願いする際に強調したことは、必要に応じた専門職員等を配置することです（図8）。研究者は研究をすることが使命ですから、人のためにデータを使えるようにして出してあげることは研究者の使命の一部として認識されていないことが多々あります。そうすると、「データを共有してください、それがモラルです、道徳です」と言っても研究者は動かないので、そこの作業をする人をきちんと手当てすることが一番大事だと考えました。そこで本事業の中では、そのような作業をする、いわばデータのアーカイブをする専門家、データの利用を促進する専門職員を配置して、その努力をしてくださいと皆さんにお願いしています。

### 拠点機関の紹介

大阪商業大学 JGSS 研究センターは、大規模社会調査を継続して行っているところなので、ここに関しては、最初は自機関のデータを集中的に作業した後、他機関のデータの保存や公開の作業に乗り出しています（図9）。

慶應義塾大学経済研究所パネルデータ設計・解析センターは非常に重要な活動をしており、ルクセンブルク・インカム・スタディという国際比較プロジェクトに日本のデータを出しています（図10）。この機関が

#### 拠点機関の活動④

##### 4. 留意事項

必要に応じた専門職員等の配置を可能にした

- ① データ共有基盤の強化に当たり、データ・アーカイブ・マネージャー、データ・アーキビスト、データ・ライブラリアンなど、研究者のデータ利用を支援する者
- ② 研究員や専門職員など、データに付加価値をつけることやデータを国際化、高度化するために必要な者
- ③ IT技術者など、高度なデータ利用環境の整備・構築を支援する者

(図8)

実施しているデータを、先進国の重要な国際比較研究のために提供しています。他の国では政府統計が提供されていることが多いのですが、日本の場合はここが出しています。ですから、このセンターが保有しているパネルデータについて逐次、機能強化の取り組みを進めていただくと同時に、他機関の重要なパネルデータについても同様の活動をしていただくという形で進めています。

東京大学社会科学研究所附属社会調査・データアーカイブ研究センターは、比較的古手ですと活動しています(図 11)。ここは私の古巣でもあり、基本的に社会調査データを収集して共有する活動を行っています。特に諸外国のデータアーカイブでは、さまざまなデータリポジトリの認証活動や、DOI を付与する活動がかなり進んでいるのですが、日本ではなかなかできていないので、さまざまな情報共有活動、情報と情

報の連結を行い、データをファインダブルにして皆さんによりよく使っていただき、国際化を進めるということをしています。

一橋大学経済研究所は、主に政府統計の加工データやマイクロデータの利用に関する活動が熱心なところ(図 12)。特に政府統計に力を入れて、学術目的の利用あるいは英語化を進める活動をしています。

### 総合データカタログについて

このように最低限、四つの拠点機関がさまざまな活動をしています。四つの機関それぞれのウェブサイトに行ってそれぞれ検索するというのは不便ですし、利用者は一括で検索してどのようなデータがあるのか見たいものです。そのような例は諸外国でもあります。

そこで、私どもは各拠点機関から提供を受ける調査データのメタデータを自動的に収集して、一括検索で

**拠点機関1 大阪商業大学  
JGSS研究センター**

- JGSSデータを整備し、公開を希望する他機関のデータ整備を助け、データの英語化と国際統合データの構築を進め、累積およびパネルデータを整備し、関連文献の情報を収集・提供する。
  - JGSSデータ/EASSデータの整理・公開準備
  - 地点情報を含んだデータの分析環境の整備
  - 長期累積データの作成
  - 他機関のデータ寄託準備、データ整理の支援

(図 9)

**拠点機関3 東京大学 社会科学研究所附属  
社会調査・データアーカイブ研究センター**

- 国際的に認められる信頼できる日本の社会科学データアーカイブとしての地位を確立することを目指す。
  - システムの強化・改善、英語化への対応、セキュリティの充実をはかる。
  - 国際化およびオープンサイエンス時代に対応したシステム整備により、日本の人文学・社会科学研究に貢献する。
  - 新たに取り組むのは、①利用者および寄託者管理システムの強化、②セルフ・アーカイブ・システムの導入、③リモート集計システムの強化、④メタデータの再整備、⑤国際認証をめざしたポリシーとシステムの整備、の5点である。

(図 11)

**拠点機関2 慶應義塾大学 経済研究所  
パネルデータ設計・解析センター**

- 日本におけるパネルデータに関するデータアーカイブの中核的な機能を担うため、データアーカイブ機能の強化(共有化)と海外発信・連携機能の強化(国際化)を進める。
  - 2018年度は、同センターが保有しているパネルデータについて機能強化の取組を進める。
  - 同センターが継続的に実施しているパネル調査データについても、各波のデータの準備が整い次第、順次提供する。
  - 同センター以外が保有するパネルデータについても、「データアーカイブ機能の強化」と「海外発信・連携機能の強化」の取組を進める。

(図 10)

**拠点機関4 一橋大学経済研究所**

- 政府統計データやその集計・加工データの利活用を促進する総合的なシステムを構築する。
  - 長期間をカバーする政府統計データや調査票情報等メタデータのデータベース化・英語化
  - 政府統計個票の集計データや加工統計の収集・整理と公開
  - データの英語化や国内外への発信
  - 個票データの利用可能性の向上

(図 12)

きる仕組みをつくるための準備をしています。そのためメタデータの記述は、ダブリンコアが一番有名だということは理解しているのですが、社会調査データのメタデータに関しては、ダブリンコアともマッピングはされている Data Documentation Initiative (DDI) という規格を使ってカタログを作ります。それで、拠点機関のメタデータだけではなく、他の機関のメタデータも同様に集めるような形で作るつもりです。

◆  
●フロア 1 東京財団政策研究所の職員です。各拠点機関で、必要に応じた専門職員の配置を可能にしたということですが、実際に職員を既に採用された拠点機関はありますか。

●前田 各拠点機関の人事については、インフォーマルにいろいろ聞くことはあるのですが、ここでオープンに言えるほど把握していません。ただ、私の知っている限りは、きちんとしたスペシフィケーションで公募しているところがあったので、各大学がどのようなスキルを必要とするかにもよるのですが、単純に研究者を採用しているわけではないという共通認識があるのではないかと理解しています。

●フロア 2 慶應義塾大学の倉田 (SPARC Japan 運営委員会委員) と申します。調査データの現状に関する調査は私も答えさせていただいて、大変時間がかかったのですが、まさにこれのオープンデータ化ということはお考えなのでしょうか。

●前田 それは、調査の案内状段階でどのような形で皆さまにお願いしたかを確認する必要があるのですが、「私の」方針としては、テクニカルに問題がなければ、公開というよりも、ご利用いただく環境を整えた方がよいと思います。これは研究者個人を確実に特定できるので、そこをどう考えるかです。ただ、そのよう

なご意見は実は他にも 2~3 件頂いておりまして、それに関しては持ち帰ってきちんと考えて、どこかの段階で公式に、お尋ねいただいたら返答できるようにしたいと思います。

●フロア 3 総合データカタログのところでご紹介いただいた、社会調査メタデータの DDI の規格についてですが、現在、ご紹介いただいた拠点機関で作成されているメタデータは、この DDI に準拠したものなのでしょうか、あるいはダブリンコアなのでしょうか。

●前田 ここは面白い点ですが、社会調査データの保存・公開活動と、いわゆる図書館の活動は元々無関係に行われていた部分が日本ではありまして、私が理解している限り、社会調査データのメタデータは特に規格はないといえますか、社会調査の専門家がこういう情報が必要だと考えたものが列挙されています。私の知っている東京大学のケースだと、HTML のファイルの中にその情報がテキストとしてあるだけで、例えば MARC なり、ダブリンコアなり、何らかのマークアップランゲージの中にそれが記述されて、検索可能になっているのではなく、かなりローカルな仕組みになっています。

それを DDI へ持っていく段階でかなりの作業が必要になるのですが、データをアクセシブルにするというときに、現在はヨーロッパやアメリカの大手のデータアーカイブはほとんど DDI ベースのカタログに移行しつつあり、諸外国のデータ、例えば英語のメタデータのインターオペラビリティなどを考えても、恐らく今の段階では DDI に移行することが得策だと考えています。ただし、現在は日本で、東京大学や他の機関で実験的に DDI ベースのカタログを作ろうという努力はしていますが、本格稼働しているものはありません。

●フロア 4 国立情報学研究所 (NII) の武田 (SPARC Japan 運営委員会委員長) と申します。今の



点に関連して、先ほど DOI を付けるといった話にも少し言及されていたと思います。それも具体的に DOI を付ける方向で考えていらっしゃるのかとか、その辺の計画を分かる範囲で教えていただければと思います。

●前田 DOI を付ける範囲ですけれども、現在は、日本学術振興会としての総合データカタログは、メタデータのカatalogで、メタデータを相互検索した後、実際のデジタルオブジェクトであるデータファイル自体は各拠点機関に取りに行くという構想になっているのです。ですので、現段階で、DOI を付ける範囲について、明確なポリシーをまだ決めていない状態です。少なくとも一つの機関は、本事業が始まる以前から、DOI に関する計画を進めていたようです。その意味で言うと、例えば本事業の中で一貫したやり方でやるべきなのか、各拠点の機関の活動に任せておくと良くないのか、逆にいろいろな形でご助言いただければと思います。

●フロア 4 私は NII にありますが、国立情報学研究所 (NII) と科学技術振興機構 (JST) と国立国会図書館 (NDL) と物質・材料研究機構 (NIMS) でジャパンリンクセンター (JaLC) というものを運営しています。DOI を付けるときに、結局、ランディングページというページの部分を誰が責任を持ってやってくれるかによって、誰が作るかが決まってくると思います。その体制の詳細がまだよく分からなかったなので、そのあたりをジャパンリンクセンターに相談していただければ、一緒に考えますので。

●前田 分かりました、ありがとうございます。