



JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE

日本学術振興会

2020年度統計関連学会連合大会
2020年9月10日13:00-15:00 C会場

人文学・社会科学分野におけるデータインフラストラクチャー
—データカタログの整備—

[2CPM1-02]

社会科学の分野におけるメタデータの整備状況について

独立行政法人日本学術振興会

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研究員(PO)

前田幸男(東京大学大学院情報学環・教授)

伊藤伸介(中央大学経済学部・教授)

1. メタデータとコードブック
2. Data Documentation Initiative (DDI) について
3. CESSDA Data Catalog
4. 人文学・社会科学総合データカタログ
5. メタデータ・スキーマや統制語彙の長期的運用について

- メタデータの定義

情報資源を効果的に識別・記述・探索するために、その特徴を記述したデータ。一般的には、データについてのデータと定義される。ネットワーク情報資源の管理と結び付いて生まれた概念である...(中略) **必ずしも新しい概念ではない**...(後略)。

『図書館情報学用語事典』第5版, 2020.

1. メタデータとコードブック

- メタデータは、統計調査・社会調査の文脈で言えば、主に「調査の概要」などの文書と、符号表を指すことになる。
 - 調査に関わるメタデータ
 - 変数に関わるメタデータ
- 大規模調査では、「調査概要」「符号表」「データレイアウト表」等一式が冊子形式で準備されることが多かった(コードブック)。
 - 現在でもPDF版のコードブックが作られることは多い。

メタデータとデータの関連付け SPSSを例に

```
FILE HANDLE DATA /  
NAME="C:¥DATA¥07294-  
0001-Data.txt"  
LRECL=1367 .  
DATA LIST FILE=DATA /  
V1 1-4  
V2 5-8  
V3 9-10  
V4 11-12  
...  
V63 158-159  
V64 160-160  
V65 161-161  
...
```

```
VARIABLE LABELS  
V1 "ICPSR調査番号 (7294)"  
V2 "面接番号"  
...  
V64 "16(a) 今の日本の憲法は変えた方が良いと思  
いますか。それとも今のままで良いと思いま  
すか。"  
V65 "16(b) (16(a)で「変えた方が良い」と言  
った人に) 大幅に変えた方が良いと思いま  
すか。それとも多少変  
える程度で良いと思いま  
すか。"  
...  
VALUE LABELS V64 1"変える" 5 "今のま  
ま" 7"  
その他" 8"わからない" 9 "答えなし" 0"  
非該当".  
VALUE LABELS V65 1"大幅" 5 "多少" 7"  
その他"  
8"わからない" 9 "答えなし" 0"  
非該当".
```

コンピューターで処理する関係上、記述のルール・制約は明確。

調査についてのメタデータ

- メタデータの項目（エレメント）としては、例えば、
 - 研究代表者
 - 調査実施機関
 - 調査概要
 - 母集団
 - サンプルングの方法
 - 調査方法
 - 調査実施時期
 - 識別子（Digital Object Identifier など）
 - 研究助成機関
- この他、「調査票」「回答票」「依頼状」なども重要なメタデータではある。

2. Data Documentation Initiative (DDI) について

- DDIは社会科学、行動科学、そして関連領域のデータを記述するためのメタデータ・スキーマ
 - メタデータ・スキーマ:メタデータ・エレメントの定義、書き方を標準化したものの集合(田村編『情報資源組織論改訂版』樹村房, 2016)
- XMLによりタグづけすることで、メタデータの項目が持つ意味を、機械が理解できるようになる。
- 一定のルールに基づき記述することにより、システム間の相互運用が可能になる。

Census of Population and Housing, 2000 (ICPSR 13511) の例

```
- <stdyDscr>
  - <citation>
    - <titlStmnt>
      <titl>Census of Population and Housing, 2000 [United States]: Public Use Microdata
        Sample: 1-Percent Sample </titl>
      <IDNo agency="ICPSR">13511</IDNo>
      <IDNo agency="dara">http://doi.org/10.3886/ICPSR13511.v1</IDNo>
    </titlStmnt>
    - <rspStmnt>
      <AuthEnty>United States. Bureau of the Census</AuthEnty>
    </rspStmnt>
    - <prodStmnt>
      - <copyright>
        <ExtLink
          URI="https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/13511/terms
            ">ICPSR Terms of Use</ExtLink>
        </copyright>
      </prodStmnt>
    - <distStmnt>
      <distrbtr abbr="ICPSR" affiliation="Institute for Social Research, University of
        Michigan">Ann Arbor, MI: Inter-university Consortium for Political and Social
        Research</distrbtr>
      <distDate date="2020-09-04"/>
    ...
```

3. CESSDA Data Catalog

About

Mission & Vision

Consortium

CESSDA ERIC

Membership

Main Office

Working Groups

Projects

Governance

History

Members

Partners

CESSDA Countries

Members

Partners

Austria

Belgium

Croatia

Czech Republic

Denmark

Finland

France

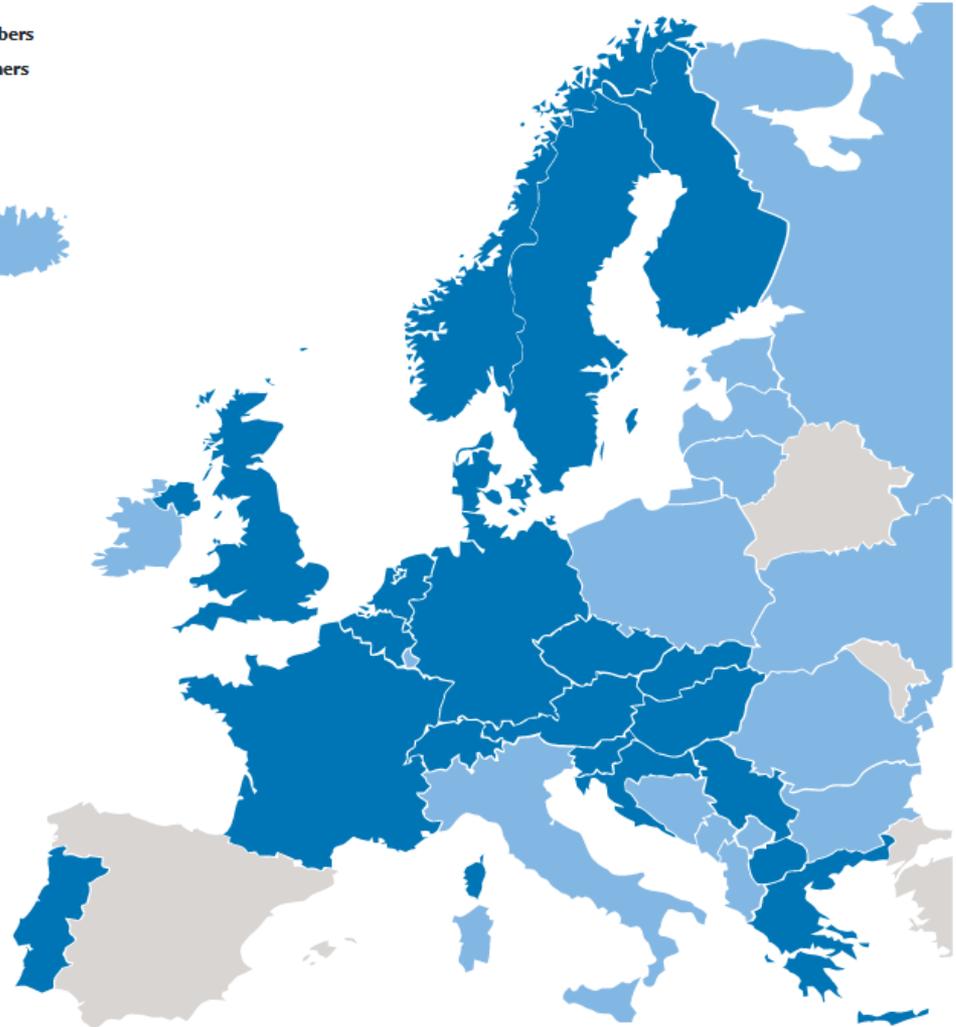
Germany

Greece

Hungary

Netherlands

North Macedonia



3. DDIに依拠したデータカタログの実例



Q "public opinion"

ヨーロッパ横断データ・カタログ

1906 results found in 229ms

Advanced search

Reset filters

Clear search

Topic ?

Results per page 30

Sort by Relevance

Collection years ?

< 1 2 3 4 ... >

Country ?

[Role of Government and Slovenian Public Opinion 1989](#)

*Tos, Niko (Public Opinion Research Centre, Faculty of Social Science, University of Ljubljana, Slovenien);
Klinar, Peter (Public Opinion Research Centre, Faculty of Social Science, University of Ljubljana, Slovenien);
Markic, B. (Public Opinion Research Centre, Faculty of Social Science, University of Ljubljana, Slovenien); (3 more)*

Functions and tasks of the state. Topics: 1. Citizen rights: law-abidingness and selected forms of protest against the government; freedom of speech and freedom of the press even for revolutionaries and racists; police intervention in the private sphere to prevent crimes; principle of rule by law, in doubt for the accused; attitude to participants in demonstrations, racists and revolutionaries. 2. Government intervention: tax progression and income re-distribution by the government; governmenten...

Publisher ?

Language of data files ?

[Occupation of Kungstorget in Göteborg 1976 - studies about information and the formation of public opinion](#)

*Kent Asp (University of Gothenburg, Department of Journalism Media and Communication);
Claes Lundgren (University of Gothenburg, Department of Political Science)*

On the initiative and with the support of The National Psychological Defense Planning Committee this study was made of medias' reporting of the events connected with the occupation of Kungstorget in Göteborg. The aim was to study information and the creation of public opinion around the occupation, which part medias' moves played for the activists' possibilities to affect a larger section of public opinion and a try to investigate what value various media have as a basis for information for t...

タイトルをクリックすると次の画面に

“Go to Publisher” をクリックするとSNDの画面に

CESSDA Catalog の画面



Q "SND_snd0115"

Advanced search

Reset filters

Clear search

Similar results

[Public opinion about social assistance in Sweden 1985](#)

[\[SJM\] Slovene Public Opinion : Series information](#)

[Quantity and Quality of Spatial Information About Public Transport in 58 European Towns, 1992-2000](#)

[Advance Factory Information System, 1976-1987; Public Sector, Grampian Region](#)

[Public Opinion about Legitimacy of the political System in Serbia 1997](#)

[Back](#)

[View JSON](#)

[Go to publisher](#)

Study title

Occupation of Kungstorget in Göteborg 1976 - studies about information and the formation of public opinion

Creator

Kent Asp (University of Gothenburg, Department of Journalism Media and Communication)

Claes Lundgren (University of Gothenburg, Department of Political Science)

Study Persistent Identifier

SND 0115 (SND)

<https://doi.org/10.5878/000065> (DataCite)

Abstract

On the initiative and with the support of The National Psychological Defense Planning Committee this study was made of medias' reporting of the events connected with the occupation of Kungstorget in Göteborg. The aim was to study information and the creation of public opinion around the occupation, which part medias' moves played for the activists' possibilities to affect a larger section of public opinion and a try to investigate what value various media have as a basis for information for the standpoint of the single citizen.

Methodology

Data collection period

Not available

Country

Sweden

Time dimension

Not available

この画面で読めるメタデータは全て次のスライドにあるスウェーデンのカタログからOAI-PMHで自動的に読み込まれたもの。

SNDのカタログの画面



Search on site | På svenska

Log in

[Find and Order Data](#) [Manage Data](#) [Describe and Share Data](#) [News & Events](#) [About Us](#) [Contact](#)

Occupation of Kungstorget in Göteborg 1976 - studies about information and the formation of public opinion

 [Overview](#)

 [Description](#)

 [Data](#)

Creator/Principal investigator(s):

Kent Asp - University of Gothenburg, Department of Journalism Media and Communication
Claes Lundgren - University of Gothenburg, Department of Political Science

Description:

On the initiative and with the support of The National Psychological Defense Planning Committee this study was made of medias´ reporting of the events connected with the occupation of Kungstorget in Göteborg. The aim was to study information and the creation of public opinion around the occupation, which part medias´ moves played for the activists´ possibilities to affect a larger section of public opinion and a try to investigate what value various media have as a basis for information for the standpoint of the single citizen.

Subject area:

media (CESSDA topic classification)
Social Sciences, Media Studies (The Swedish standard of fields of research 2011)

Keywords:

internal politics

Principal organisation:

[University of Gothenburg](#)

Accessibility status:

 [Access to data through SND](#)
Data are accessible by order

 [Order data](#)

SNDのカタログ。DDIの仕様
に従いメタデータを準備

4. 人文学・社会科学総合データカタログの計画

- 拠点機関が収集・整理・提供するデータを横断的に一括検索できるデータカタログを構築する。
- 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OAI-PMHを通じて自動的に収集する。
- CESSDA, ICPSR, ADA, UKDS, GESIS, FSD等のカタログを参考にメタデータ・スキーマを設計
 - 海外機関がファセット(検索フィルター)として採用しているエレメントに特に着目
 - 素案をもとに海外機関に助言を求める
- 最終的には31の項目を採用
 - 項目を 必須、強く推奨、推奨、任意 の4つに分類

各国カタログの比較作業(抜粋)

CESSDA	UKDA	GESIS	DANS	ADA
Study title	Title:	Title	Persistent identifier	Title
Creator	Study number (SN):	Abstract	Title	Dataset Persistent ID
Study Persistent Identifier	Persistent identifier:	Data Version	Creator	Publication Date
Abstract	Series:	Keywords	Contributor	Other ID
Country	Principal investigator(s):	Processing	Date created (ISO 8601)	Alternative Title
Time dimension	Depositor:	Geographical Coverage	Description	Alternative URL
Analysis unit	Sponsor(s):	Organizaiton	Audience	Author
Sampling procedure	Grant number	Universe	Subject	Contact
Data collection method	First edition release	Sampling	Temporal coverage	Description
Data collection period	Latest edition release	Data Collection	Spatial coverage	Subject
Language of data files	Topics	Anonymization	Identifier	Keyword
Publisher	Abstract	Legal Basis	Relation	Topic Classification
Year of publication	Dates of fieldwork:	Weighting	Type (DCMI resource type)	Related Publication
Terms of data access	Country:	Data Access	Format	Note
Study number	Spatial units:	Access Conditions	Language (ISO 639)	Language
Topics	Observation units:	Access Form	Language	Producer
Keywords	Observation unit location:	Access Contact	Source	Production Date
	Population:	Data Service	Remarks	Production Place
	Number of units:	Comparability	Access rights	Contributor
	Method of data collection:	Comparability over time	Date submitted	Distributor
	Time dimensions:	Comparability between Countries		Distribution Date
	Sampling procedures:	Note		Depositor
	Kind of data:	References		Deposit Date
	Weighting:			Time Period Covered
				Date of Collection

統制語彙について(1)

- 情報検索において、索引語として利用する語を限定し、その意味範囲や使用方法を規定したもの…基本的な目的は、語形や表記の変化、同義語の存在で検索漏れが生じないように…すること…

『図書館情報学用語事典』第5版, 2020.

- ファセットによる分類のためにも重要。
 - 調査地域や言語については既存の統制語彙 (ISO) を借用
 - アクセス権やデータの分類については独自の分類 (語彙) を作成

統制語彙について(2)

- 社会調査に関わるものはDDI Alliance が提供している統制語彙から3つを翻訳
 - 観察単位(分析単位)
 - サンプルング方法
 - 調査方法
- Topic については複数の語彙を併用
 - 海外からの利用を念頭にCESSDA Topic Classificationを翻訳
 - 公的統計の統計表を念頭に日本統計年鑑目次を利用

5. 長期的運用について

- データが複雑化するに従い、データを記述するメタデータのルールも変化する。
- 社会状況が変化するに従い、研究対象を表す語彙も変化する。
- 研究の基盤を維持するためには、メンテナンスやアップデートが必須。
 - データ・ファイル自体のアップデート
 - 機器やプログラムのアップデート
 - メタデータ・スキーマや統制語彙のアップデート