Initiatives for Sharing and Utilization of Humanities and Social Science Data at the University of Tokyo

東京大学における人文学・社会科学データの 共有・利活用に関する取組

Yamada, Taizo Yokouchi, Nobutada Nakamura, Satoru



Miwa, Satoshi

Shibutani, Ayako

Hirasawa, Kanako



ISS INSTITUTE OF SOCIAL SCIENCE, THE UNIVERSITY OF TOKYO

Project



Program for Strengthening Data Infrastructure for the Humanities and Social Sciences

- -By Japan Society for the Promotion of Science
- -From 2023 Until 2027 (tentative)
- -(For short "Current Program")
- -This is a program premised on the Program for Constructing Data Infrastructure for the Humanities and Social Sciences, 2018-2022.
 - (For short "Constructing Program")

Objective

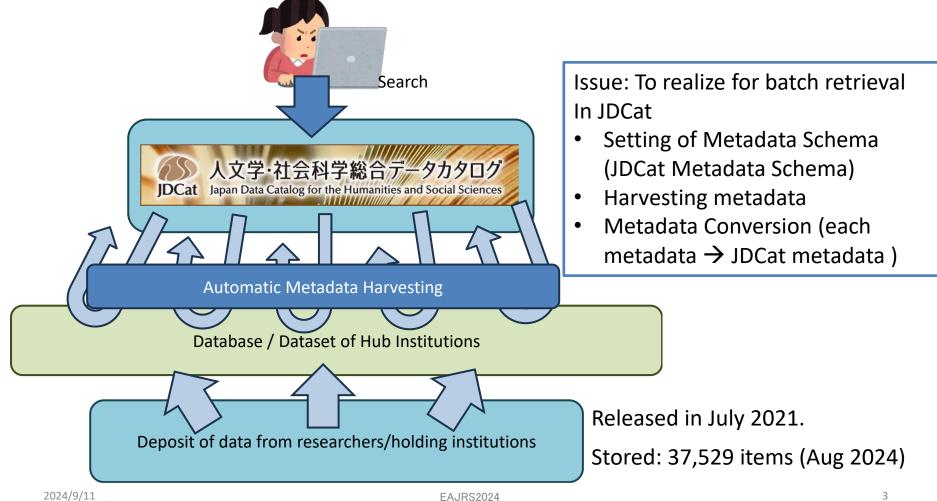
- -Enhancement and reinforcement of data platforms and infrastructures to promote data sharing and utilization in the humanities and social sciences
- -Aiming to encourage researchers to share data with each other and to promote joint research, etc. in Japan and other countries.

JDCat Japan Data Catalog for the Humanities and Social Sciences



JDCat was created to allow the batch retrieval of metadata (e.g. the contents, contexts and sources of data sets) across fields of the humanities and social sciences.

Metadata from the research institutions commissioned to operate this program ("Hub institutions") can be retrieved in batches.



JDCat

https://jdcat.jsps.go.jp



Language: English ~

Japan Data Cata for the Humanities and Social Sci			
Top Information About Manual Li	ık Guideline JDCat Analysis Tool		
	Type and press enter to search	Q Search V Advanced	
	Full text C K	Keyword	
Topic 🔹	Item Lists	1 - 20 of 2254 res	ults.
Select 🗸 🗸		Display Order: Title V asc V Display Number 20	~
	▶ (阿部正弘書状)・(守玄院一ニツキ上申書)・嘉祥ニ付	対登城之書付・刑部卿殿逝去ニ付御機嫌伺十五日より廿日迄之日記手留一件_6	
Distributor Institute of Economic Research, Hitotsubashi U niversity(34293)	阿部, 正弘 Historiographical Institute, the University of Tokyo materials: text		
Historiographical Institute, the University of Tok yo(2254)	(阿部正弘書状)・(守玄院一ニツキ上申書)・嘉祥二付	寸登城之書付・刑部卿殿逝去ニ付御機嫌伺十五日より廿日迄之日記手留一件_2	
 SSJDA(867) JGSS Research Center(81) The Panel Data Research Center, Institute for E conomic Studies, Keio University(32) 	阿部, 正弘 Historiographical Institute, the University of Tokyo materials: text		
Data Language	(阿部正弘書状)・(守玄院一ニツキ上申書)・嘉祥二付	対登城之書付・刑部卿殿逝去ニ付御機嫌伺十五日より廿日迄之日記手留一件_5	
_ jpn(37523) _ eng(123)	阿部, 正弘 Historiographical Institute, the University of Tokyo materials: text		
	■ (阿部正弘書状)・(守玄院一ニツキ上申書)・嘉祥二付	寸登城之書付・刑部卿殿逝去ニ付御機嫌伺十五日より廿日迄之日記手留一件_3	

https://jdcat.jsps.go.jp





ISS INSTITUTE OF SOCIAL SCIENC THE UNIVERSITY OF TOKYO

Metadata (1)

JDCat Metadata Schema

Providing elements that could be set as metadata related to the humanities and social sciences.

Assignment:

- DOI (Document Object Identifier)
- Access right
- License

Faceted items in JDCat: Prepared as a controlled vocabulary.

No.	JDCat element	JDCat エレメント名(日本語)	記入レベル (社会科学) /Usage(Socia I Science)	記入レベル (人文学) /Usage(Hu manities)	同一言語での繰 り返しの可否 /Repeatable	統制語彙の 有無 /Controlled Vocabularie s	ファセット 対象 /Faceted Search
1	l Title	タイトル	O	O	×		í
2	2 Study ID	整理番号	0	0	×	1	
3	3 Author	作成者	0	0	0		
4	1 Distributor	配布者	0	0	×		0
5	5 URI	URI	0	0	×		
6	5 Topic	トピック	0	0	0	0	0
7	7 Summary	概要	0	0	×		
8	3 Time Period(s)	対象時期	0	0	0		0
9	9 Geographic Coverage	対象地域	0	0	0	0	0
10) Data Type	データタイプ	0	0	0	0	0
11	l Access	アクセス権	0	0	×	0	0
12	2 License	権利情報	0	0	×	-	[
13	3 DOI	DOI	0	0	×		
14	1 Data Language	データの言語	0	0	0	0	0
15	5 Version	バージョン情報	0		×	-	
16	0 Unit of Analysis	観察単位	0		0	0	
17	7 Universe / Population	母集団	Δ		×		
18	3 Sampling Procedure	サンプリング方法	Δ		0	0	[
19	Collection method	調査方法	Δ		0	0	
20) Funding Agency	研究助成機関			0	-	[
21	l Grant ID	研究費番号			0		
22	2 Alternative title	その他のタイトル			0		ſ
23	3 Series	シリーズ	******		×	4.0000000000000000000000000000000000000	
24	Date of collection	調査日			0		
25	5 Sampling Rate	回収率			×		
26	Bibliographic Citation	引用上の注意			×		
27	7 Datafile URI	データファイルURI	******		0		
28	Related Studies	関連情報			0		
29	Related Publications	関連文献	******		0		
30	Publisher	編集者			×		
31	l Provider	所蔵者・寄託者	******		×		ſ



SS INSTITUTE OF SOCIAL THE UNIVERSITY OF T

Controlled vocabularies



Element	CV				
Торіс	CESSDA Topic Classification (Version 4.2) Japan Statistical Yearbook Classification Nippon Decimal Classification Distinction Table				
Geographical coverage	ISO 3166 Country Codes Japan Statistical Yearbook Old Geographic Coverage				
Data type	Original				
Access	JPCOAR				
Data language	ISO 639 Code Tables Codes for the representation of names of scripts				
Unit of analysis	DDI Analysis Unit (Version 1.0)				
Sampling procedure	DDI Sampling Procedure (Version 1.1)				
Collection method	DDI Mode Of Collection (Version 3.0)				
Funding agency	Original				
https://jdcat.jsps.go.jp/JDCatmetadata.html					

Metadata (2)



For batch retrieval: JDCat metadata

-We have separately prepared metadata to be provided to users and metadata to be harvested from hubs.

JPCOAR : For Humanities

- -JPCOAR schema is aimed at improving the international interoperability of Japanese institutional repositories and promoting the smooth distribution of Japanese research products.
- -https://schema.irdb.nii.ac.jp/en

DDI (Data Documentation Initiative): For Social Science

- DDI is a metadata schema for describing data in the social and behavioral sciences and related fields.
- -DDI is used in the CESSDA Data Catalog (\rightarrow Standard Metadata)
- -https://ddialliance.org/

Metadata Harvesting



Two Types available:

-OAI-PMH(Open Archives Initiative Protocol for Metadata Harvesting)

-ResourceSync

Harvest frequency: once a month

-From each institution's repository

Data, Dataset (1)



Materials of Bingo Fukuyama Abe Family(備後福山阿部家史料)

The collection consists of 776 items, including letters and diaries.

:= 隠す	── 東京大学 SHIPS Image Viewer	所員用ログイン
▲ 4 / 63 ▶	、 全画面 ◎ホーム ●拡大 ● 総小 ○ 左回転 ● 印刷ブレビュー ##一覧へ	
₩ ₩-10 +10 ₩		
00000001		
0000002	2. 44 X W 03 20, 17	
0000003		E
00000004	了动生的苦白花 七	£ .
00000005	高九型 人民的 百	54 & 15 4049
0000006	唐天町 登程展	
0000007	is the state of the second sec	12 303
0000008	A 2 12/5	AN MA
0000009	等 就 奉考室	
00000010	the stars	S.S.
00000011	1、小学生 西和	to I'm III
00000012	F & Fr when the	a the first
00000013	一根 司 雷星 人名休	
00000014		2 2 2 A
00000015	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	hart site
データベース名 : 所蔵史料目録データベ ース (Hi-CAT) 書 名 : 備忘録(寺社奉行日記)	要"例》希子	ACADO ANS
画像情報	20 2243	
Manifest Universal Viewer Mirador	PermaLink https://clioimg.hi.u-tokyo.ac.jp/viewer/image/idata/T50/07/008/00000004.jpg	

Held by Historiographical Institute, the University of Tokyo

Historical material on the Abe family in the 18th and 19th centuries. The head of the Abe family held important positions in the Edo Shogunate.

Important historical material for the study of the Edo Shogunate.

Data, Dataset (2)

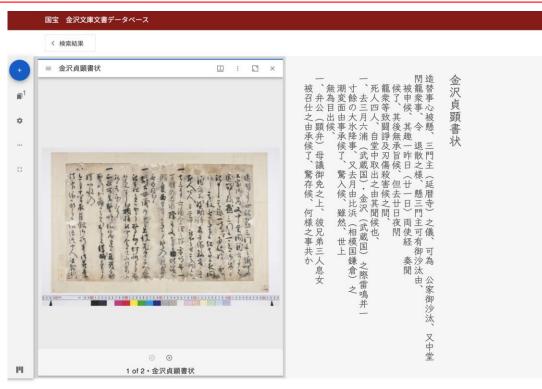
愛京大家史料錫集所 HISTORIOGRAPHICAL INSTITUTE, The University of Tokyo そISS INSTITUTE OF SOCIAL SCIENCE, THE UNIVERSITY OF TOKYO

The Kanazawa Bunko Documents (金沢文庫文書)

The collection consists of 4,149 items. (Mainly letter)

Held by Kanagawa Prefectural Kanazawa-Bunko Museum

The Kanazawa Bunko Documents are designated as a Japanese National Treasure. Fundamental historical material for the study of medieval Japanese history It is also notable as a document that conveys the situation in East Asia from the late 13th to the early 14th century, which was greatly influenced by the Mongol Empire.



金沢貞顕書状

金文番号	14					
和暦年月日	年未詳					
成立	鎌倉時代後期					
員数	1通					
形状	竪紙 (礼紙)					
料紙	楮紙					
法量 縦x横	縦34.1 x 横55.2					
紙数	一紙					
紙背	題未詳聖教(綴葉装)					
付加情報	向山 致利等。					

Data, Dataset (3)



Center for Social Research and Data Archive

tute of Social Science. The University of Toky

SSJDA:

The Social Science Japan Data Archive

SSJDA collects, maintains, and provides access to the academic community, a vast archive of social science data (quantitative data obtained from social surveys) for secondary analyses.

Provided by Institute of Social Science , the University of Tokyo

Right data set:

"Survey on Changes among Women:" Women's Life Today and in the Future, 1980 「女性の変化調査」女性の生活の現在と将 来, 1980

Results of a joint study of women in order to grasp, from the most comprehensive perspective possible, "changes in the social environment surrounding women" and "the actual state and consciousness of women in the fields of life and work."

	Abstract
Survey Number	0034
Survey Title	"Survey on Changes among Women." Women's Life Today and in the Future, 1980
Depositor	Japan Institute of Life Insurance
	For detailed information, please refer to For Data Users' on the SSIDA website.
Restriction of Use	- Apply to SSJDA. SSJDA's approval is required.
Educational Purpose	Available for both research and instructional purposes.
Period of Data Use Permission	Available for both research and instructional purposes.
Access to Datasets	Download
SSJDA Data Analysis	Not available
Summary	The Japan Institute of Life Insurance and Nomura Research Institute conducted a joint study of women in order to grasp, from the most comprehensive perspective possible. "changes in the social environment surrounding women" and "the actual state and consciousness of women in the fields of life and work." As part of the study, two questionnaires are introduced here: one directed to citizens at large and the other, to businesspersons. The study also includes interviews with and a questionnaire survey of well-informed popele as well as an analysis of data. The citizen' questionnaire are inst a clarify women's life structure, values, and work consciousness, and its terms include! Ge structure, values, and work consciousness, and its terms include! Ge structure, values, and work consciousness, and its terms include! Ge structure, values, and work consciousness, and its terms include! Ge structure, values, and and work consciousness, and its terms include! Ge structure, values, and exitive work consciousness, and its terms include! Ge structure, values, and work consciousness, and its terms include! Ge structure, values, and work consciousness, and its terms include! Ge structure, values, and work consciousness, and its terms include! Ge structure, values, and and work consciousness, and its terms include! Ge structure, values, and the related to the use consciousnes of work, and activities outside the household. The purpose of the business people's questionnaire is to investigate women's changes is a beeved by a sector-manging class of the corporate management, and it looks into changes in selected attributes of employed women, changes in the form of employment and job description, changes in values consciousness, and other related matters.
Data Type	quantitative research: micro data
Universe	Businesspersons' questionnaire 2500 managers at section management levels at listed companies nationwide, and 500 personnel department managers at listed companies nationwide. 2 canarci atticking questionnaire Women aged between 18 and 59.
Unit of Observation	Individual,Organization
Sample Size	1.581 responses from a sample of 3,000 managers (Rate of return: 20.6%) 2.2,423 responses from a sample of 3,000 women (Rate of return: 80.8%)
Date of Collection	1980-07 - 1980-10 1. July - August, 1980 2. October, 1990
Time Period	1980 - 1980
Spatial Unit	Japan 2. All over Japan (220 localities)
Sampling Procedure	Probability: Stratified Probability: Multistage Other 1. Section managers born in 1935 and after were randomly sampled from the Diamond directory of company employees. (Note that the sampling ratio was adjusted to avoid a wide inter-company variation in the number of individuals sampled.) Personnel department managers were randomly sampled (one per company) from listed companies. (Note that their representativeness of the population is limited due partly to unavailability of lists.) 2. Stratified two-stage random sampling
Mode of Data Collection	Face-to-face interview Self-administered questionaire: Paper 1. Record and mail systems 2. Record and interview systems
Investigator	Japan Institute of Life Insurance and Nomura Research Institute
DOI	10.34500/SSJDA.0034
Sponsors (Funds)	
Related Publications (by the Investigator)	Japan Institute of Life Insurance (September, 1981) "Josei no henka chosa: josei no seikatsu no genjo to shorai-dai-III-ki wo chushin toshite [The "Survey of Women's Changes" The Present and Future of Women's Life: Centering on Phase III]
Related Publications (based on Secondary Analysis)	List of related oublications (based on Secondary Analysis)
Documentation	[Kigyojin Chosa-hyo][Jopan shimin Chosa-hyo]
Maior Survey Items	 For businesspersons For businesspersons<

SS1n

Data, Dataset (4)



LTES Database:

Long-Term Economic Statistics

Database of the "Long-Term Economic Statistics" (Kazushi Ohkawa, Miyohei Shinohara, and Mataji Umemura, Toyo Keizai Shinposha, Inc., 1965-1988) which are a systematic collection of estimated and processed historical statistics of early modern Japan on economic activities in various fields based on the System of National Accounts.

Provided by Institute of Economic Research, Hitotsubashi University

Right data set:

Per Capita Personal Consumption Expenditures in 1934~1936 Prices, 1874~1940: Estimates of longterm economic statistics of Japan Personal Consumption Expenditures Table 4 1人当り実質個人消費支出(1874~1940):長期経 済統計 個人消費支出 第4表

Food, Clothing, Housing, Fuel & Light, Medical & Personal Care, Transportation, Communication, Social Expenses, Education Recreation & Misc., Total

						各,1874~				
	Table 4	Per Capita	Personal C	Consumptic	on Expendit	ures in 193	84∼1936 P	rices, 1874	1~1940	
	48323	48324	48325	48326	48327	48328	48329	48330	48331	48332
	1	1	1	1	1	1	1	1	1	1
	一人当り舅	一人当り身	一人当り身	一人当り身	一人当り到	一人当り身	一人当り身	一人当り身	一人当り身	一人当り
	Per Capita	Per Capita	Per Capita	Per Capita	Per Capita	Per Capit				
	J060400	J060400	J060400	J060400	J060400	J060400	J0604_00	J0604_00	J060400	J0604_0
	1874	1874	1874	1874	1874	1874	1874	1874	1874	1874
	1940	1940	1940	1940	1940	1940	1940	1940	1940	1940
年	円:R	円:R	円:R	円:R	円:R	円:R	円:R	円:R	円:R	円:R
Year	in yen	in yen	in yen	in yen	in yen	in yen				
1874	52.41	2.32	10.63	2.29	1.74	0.11	0	9.92	2.17	81.60
1875	54.41	2.28	11.06	2.28	1.88	0.12	0.00	9.99	2.40	84.41
1876	55.12	1.80	11.27	2.27	2.03	0.13	0.01	9.88	2.04	84.55
1877	56.39	2.25	11.26	2.29	2.08	0.14	0.03	9.91	2.18	86.53
1878	55.23	2.83	11.19	2.28	2.15	0.15	0.04	9.83	2.18	85.89
1879	57.93	3.68	11.21	2.29	2.26	0.17	0.05	9.93	2.63	90.15
1880	62.85	3.44	11.31	2.30	2.38	0.18	0.05	9.87	3.26	95.65
1881	59.07	2.68	11.15	2.36	2.50	0.19	0.05	9.84	3.38	91.22
1882	60.22	2.34	11.04	2.39	2.66	0.20	0.06	9.85	2.88	91.64
1883	58.36	2.46	10.87	2.41	2.81	0.20	0.07	9.83	2.63	89.64
1884	62.64	2.32	11.09	2.44	3.08	0.20	0.09	9.98	2.45	94.29
1885	58.55	2.82	11.25	2.47	3.33	0.19	0.06	7.82	2.51	89.01
1886	62.38	3.50	11.15	2.49	3.28	0.23	0.08	7.47	2.94	93.51
1887	65.78	3.43	11.32	2.50	3.39	0.26	0.09	9.41	4.53	100.71
1888	65.42	4.22	11.46	2.50	3.54	0.32	0.09	8.71	5.17	101.43
1889	67.48	4.63	11.45	2.56	3.71	0.38	0.10	11.16	5.59	107.07

Organization



■Core Institution: The University of Tokyo(東京大学)

- -Historiographical Institute (史料編纂所)
 - Japanese historical materials, such as Materials of the Bingo Fukuyama Abe family.
- -Institute of Social Science(社会科学研究所)
 - data on labor surveys and social surveys (in SSJDA).

Hub Institution:

- -Kobe University Library (神戸大学附属図書館)
 - newspaper articles from about 150 years ago in Japan, data related to the Great Hanshin-Awaji Earthquake.
- Nara National Research Institute for Cultural Properties (奈良文化財研究所)
 - data such as reports on archaeological sites in Japan.

System Management Institution (JDCat construction and management):

--The National Institute of Informatics(国立情報学研究所; NII)



- Collaborations with JDCat
- Accepting data from researchers and other institutions
- Promotion of data utilization
- Cooperation with core institution

Core institution are also required to function as hub institutions

Data, Dataset(5) 神戸大学附属図書館 震災文庫 https://da.lib.kobe-u.ac.jp/da/eqb/?lang=1





Earthquake Disaster Materials Collection Digital Gallery



Yoshimichi Ogimoto (5880)

ここに公開する約22,000枚の写真は、放送カメラマンの大木本美通氏がお仕事のかたわら個人的に撮影さ れたものです。

震災直後から2002年1月まで、大木本さんのお仕事場に近い三宮・元町周辺のほか、神戸市内の様々な地 域を丹念に歩かれて、この膨大な資料が生まれました。

今回、大木本さんの記憶・記録と震災文庫による調査を総合して、できる限り撮影場所の特定・キャプショ ン付与を行い、貴重な記録写真をインターネットを通じて広く公開します。

撮影者、大木本美通さんのメッセージ

6.400名余の犠牲者を出した阪神・淡路大震災から7年と5ヶ月が経ちました。 倒壊した建物や道路、 鉄道、港湾施設、廃墟と化した火災被害の商店街や住宅地... 震災後、神戸の被災地を歩き始めて感じたこ と、それは人々の生活再建に対する不安感、総てを失った喪失感であったように思います。

しかし反面、"一日も早く元の生活を取り戻したい""負けてたまるか"という被災者の強い意志の表れは、全国 から寄せられた救援の手やボランティアの活動、地域住民同志の支え合い、行政側の復興施策の後押しもあ って、厳しさの中にも着実に再興への歩みを刻んでいるように思います。

写真は傷ついた街の復旧と復興、慰霊と鎮魂、防災と新しい街づくり、他に神戸まつりや明石海峡大橋開 業、淡路花博など復興期と並行して開催された催しも含め折々に写したものを掲載しました。

記憶を記録にとの思いで写し始めたのがきっかけでしたが、この度、神戸大学附属図書館震災文庫のお力 添えで公開される機会を得られましたこと、震災の教訓を生かすうえで、今後多少なりともお役にたつこと ができれば、と思っております。

2002年6月10日 大木本美通(撮影者)

https://lib.kobe-u.ac.jp/da/about

Data, Dataset(5) 奈良文化財研究所 全国遺跡報告総覧 https://sitereports.nabunken.go.jp/en





EAJRS2024

Function as a core institution



- The Core institution is required to carry out the following tasks:
 - -Operation and modification of JDCat, and ongoing examination of the appropriateness of metadata schema and control vocabulary.
 - -Coordination with the hub institutions.

-Promotion of JDCat and this program

Web site

https://www.hi.u-tokyo.ac.jp/di/en



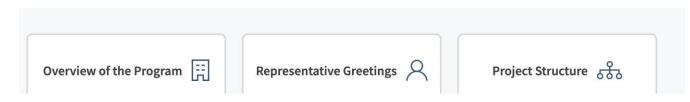
Program for Strengthening Data Infrastructure for the Humanities and Social Sciences Historiographical Institute and Institute of Social Science. The University of Tokyo / Participation Institution (Headquarters) for the USPS Rommissioned Program



Overview of the Program

The "Program for Strengthening Data Infrastructure for the Humanities and Social Sciences" aims to enhance and strengthen JDCat. This data platform promotes data sharing and utilization in the humanities and social sciences so that researchers can share data and promote collaborative research in Japan and abroad.

This website is operated by the Historiographical Institute and the Institute of Social Science, The University of Tokyo, commissioned by the JSPS as the Participating Institution (Headquarters) of the JSPS "Program for Strengthening Data Infrastructure in the Humanities and Social Sciences" Grant Number JPJS00320231001.



Conclusion, Future work



- "Program for Strengthening Data Infrastructure for the Humanities and Social Sciences"
 - JDCat, metadata, contents, organization and its functions
- The development of research data infrastructures is progressing in many countries, For example, in Europe,
 - DARIAH is a leading initiative for Digital Humanities.
 - SESSDA (Consortium of European Social Science Data Archives) is an initiative to realize cross-search of social science data archives.
- This program is the only research data infrastructure for the humanities and social sciences in Japan.
 - Irreplaceable.
 - However, the budget is very small.

Future work

- There is no crosspoint between data from the humanities and social sciences yet.
 - It is necessary to improve the environment so that collaborative research that intersects data from the humanities and social sciences can be realized.
 - For this purpose, it is essential to increase the data in JDCat. The addition of data should be promoted.
- It is essential to provide data that can be used for AI and machine learning.
 - Currently, metadata can be downloaded from JPCOAR and DDI.
 - Links to the actual contents are provided by DOI...
 - We would like to re-examine the method of use considering the diversity of the data.



Thank you for listening to my presentation.

This work was supported by JSPS KAKENHI Grant Number JPJS00320231001 (Program for Strengthening Data Infrastructure for the Humanities and Social Sciences).

24H00011(Grant-in-Aid for Scientific Research(S)), and 23HP7002(Grant-in-Aid for Publication of Scientific Research Results(Database)), too.