JOSS2026　「データ共有がひらく社会－データ提供機関における課題と展望」

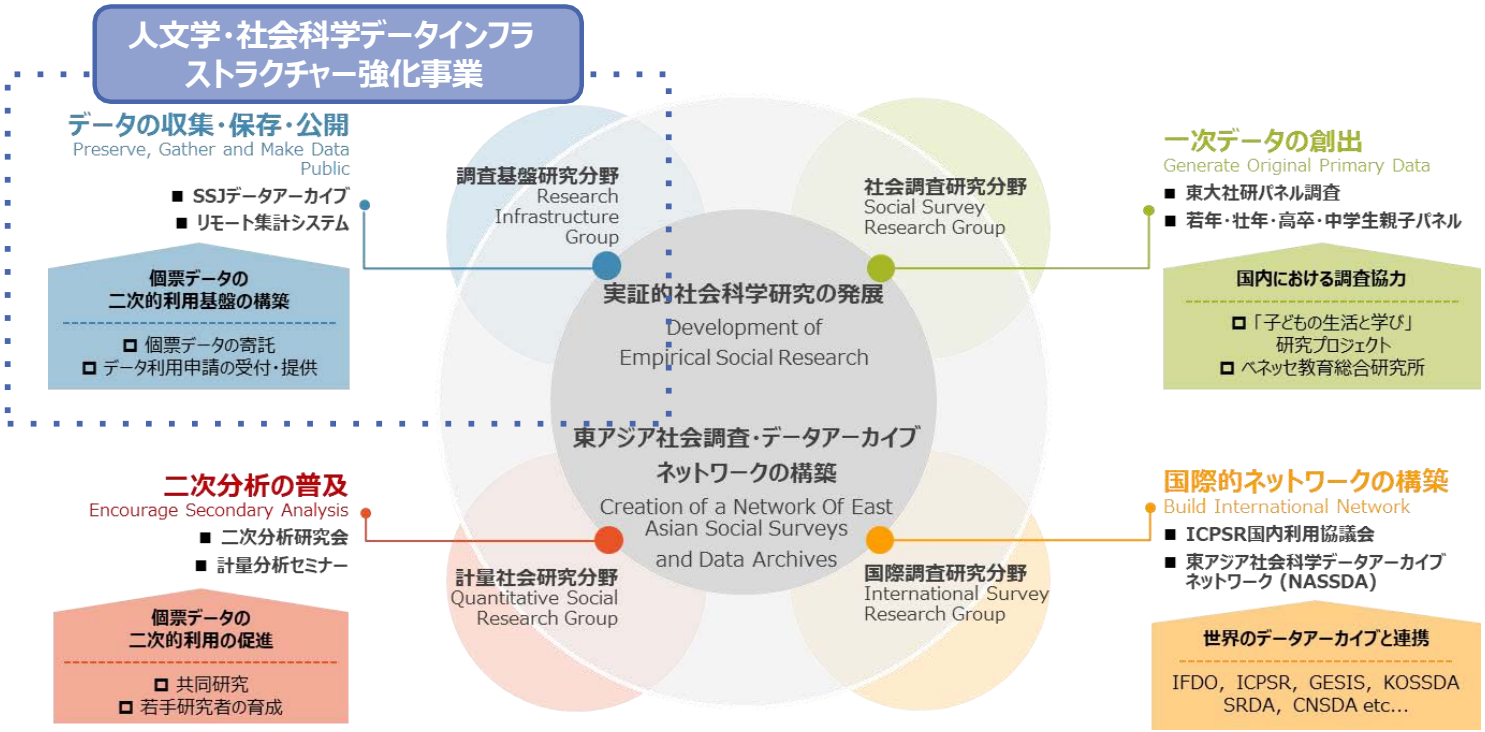
SSJデータアーカイブにおける取り組み：これまでの成果と課題、今後の展望

2026年６月２４日

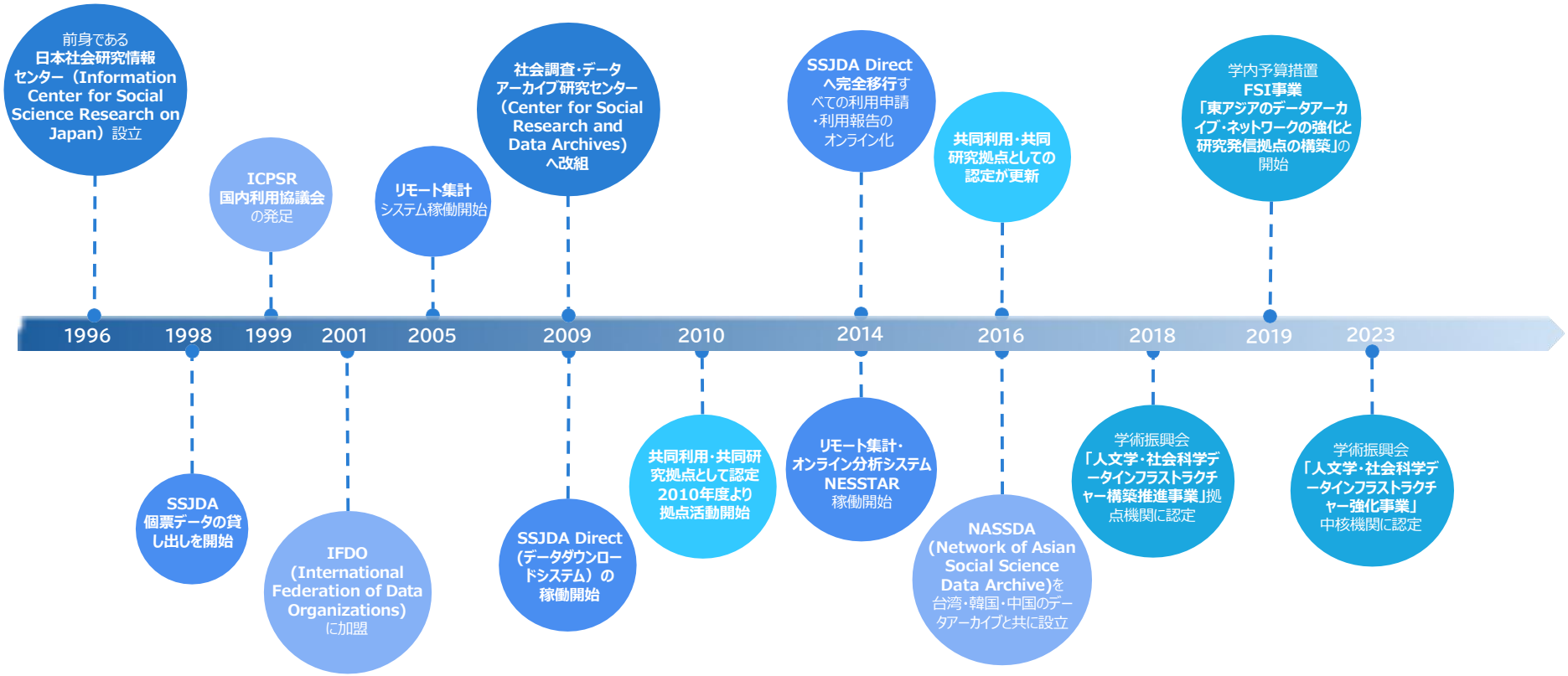
田中　慶子
東京大学社会科学研究所
附属社会調査・データアーカイブ研究センター

SSJデータアーカイブとは

- 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 (CSRDA) の事業のひとつ
  - SSJデータアーカイブ (Social Science Japan Data Archive: SSJDA)



社会調査・データアーカイブ研究センター（CSRDA）の沿革



SSJデータアーカイブの役割・意義・特長

- 役割
  - データの収集・保管と共有

質問紙

問7.　現在（または直近）のお仕事に関して、次にあげるＡからＧのそれぞれについて、あてはまる程度をお答えください。（○はそれぞれにつき１つ）

	かなりあてはまる	ある程度あてはまる	あまりあてはまらない	あてはまらない	部下はいない
A. 自分の仕事のペースを、自分で決めたり変えたりすることができる	1	2	3	4	
B. 職場の仕事のやり方を、自分で決めたり変えたりすることができる	1	2	3	4	
C. 部下の仕事のやり方を、自分が決めている	1	2	3	4	5
D. 教育訓練を受ける機会がある	1	2	3	4	
E. 仕事を通じて職業能力を高める機会がある	1	2	3	4	
F. 子育て・家事・勉強など自分の生活の必要にあわせて、時間を短くしたり休みを取るなど、仕事を調整しやすい職場である	1	2	3	4	
G. 今後１年間に失業（倒産を含む）をする可能性がある	1	2	3	4	

あなたのふだんの生活についておうかがいします。

問8.　あなたはどのくらいの頻度で以下のことをしていますか。（○はそれぞれにつき１つ）

	毎日	週に5～6日	週に3～4日	週に1～2日	月に1～3日	ほとんどしない
A. 運動（ウォーキング・ジョギング・エアロビクス・水泳・テニスなど）	1	2	3	4	5	6
B. １日に３食を食べる	1	2	3	4	5	6
C. 栄養バランスの取れた食事を取る	1	2	3	4	5	6
D. カップ麺やファーストフードを食べる	1	2	3	4	5	6
E. 食事の用意	1	2	3	4	5	6
F. 洗濯	1	2	3	4	5	6
G. 家の掃除	1	2	3	4	5	6
H. 日用品・食料品の買い物	1	2	3	4	5	6
I. 友人・恋人（配偶者は除く）と食事をする	1	2	3	4	5	6
J. 友人・恋人（配偶者は除く）と話をする	1	2	3	4	5	6
K. インターネットを利用する（仕事以外で）	1	2	3	4	5	6

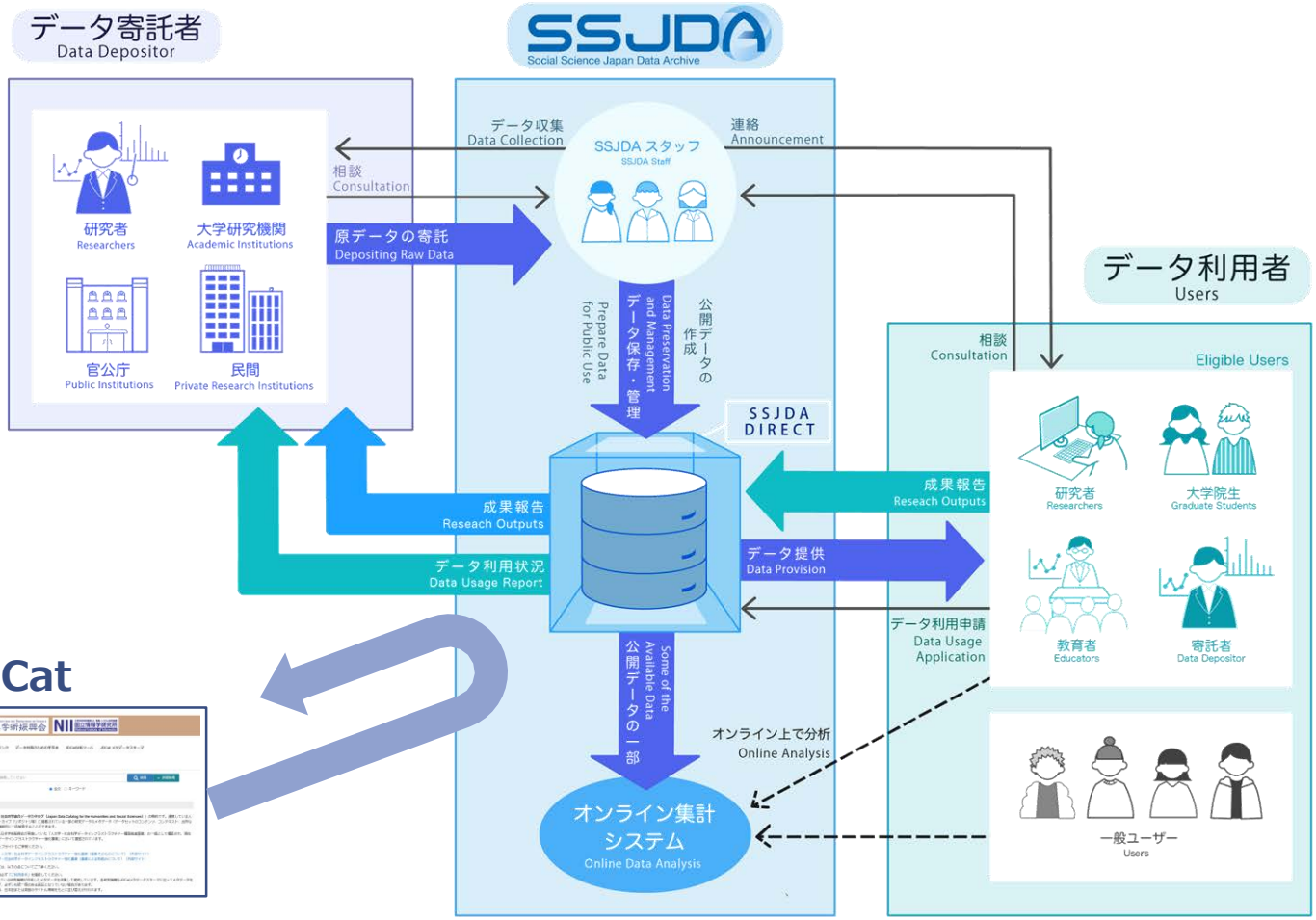
個票データ

	ZQ07A	ZQ07B	ZQ07C	ZQ07D	ZQ07E	ZQ07F	ZQ07G
1	2	3	1	2	2	4	4
2	2	2	2	4	1	2	3
3	2	3	3	2	4	2	4
4	1	2	2	2	1	2	4
5	3	2	3	4	3	1	4
6	2	2	2	3	2	4	3
7	2	3	4	2	2	2	3
8	2	4	5	4	1	1	4
9	2	3	4	1	2	2	3
10	2	2	5	3	2	2	4
11	2	2	5	3	1	3	3
12	3	2	2	2	2	3	1
13	2	4	4	4	3	3	4
14	2	3	4	2	3	4	4
15	1	3	5	2	2	1	4

SSJデータアーカイブの役割・意義・特長

- 役割
  - データの収集・保管と共有
- 意義
  - データの散逸を防ぎ，安心して長期保存できる
  - データが見つけやすく，アクセスしやすく，相互運用しやすく，再利用しやすくなる
  - 新たな視角からの二次分析が可能になったり，分析の再現性を担保できる
  - 大学院生や若手研究者にとっての研究資源であり，学部生に対する教育資源にもなる
- 特長
  - 幅広い機関・研究者の方々から調査データが寄託されている
  - 公開データを使った研究や教育を支援している
  - データアーカイブ間の学際的，国際的な連携を行っている
  - 寄託者や利用者への親身なサポートを行っている

SSJデータアーカイブの仕組み



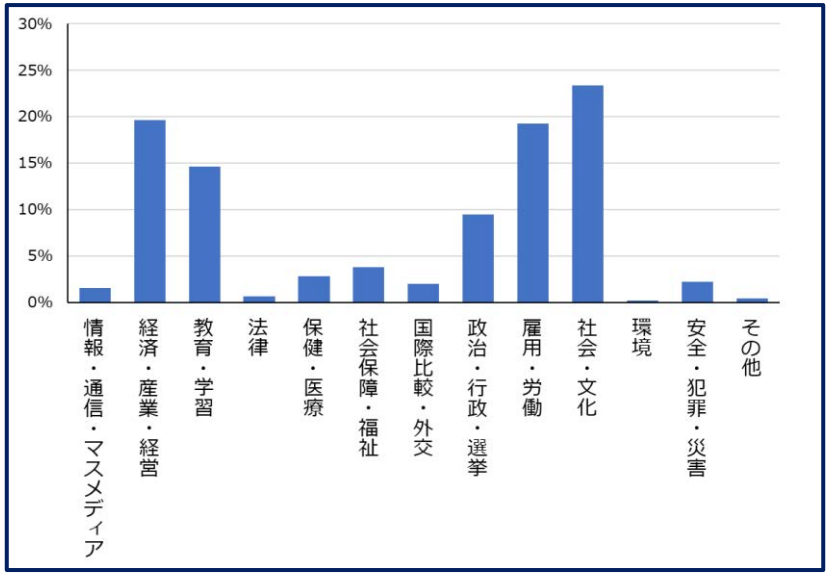
データ寄託・公開の促進

- 幅広い組織・研究者からのデータ寄託
  - これまでに200以上の組織・研究者から寄託
  - 約1,900の調査データを公開・提供

代表的な調査データと寄託者

JLPS（働き方とライフスタイルの変化に関する全国調査）
SSM（社会階層と社会移動全国調査）
JES（投票行動の全国的・時系列的調査研究）
全国就業実態パネル調査（リクルートワークス研究所）
学校教育に対する保護者の意識調査（ベネッセ教育総合研究所）
日本人の意識調査（NHK放送文化研究所）
食育に関する意識調査（農林水産省）
仕事と育児の両立に関する実態把握のための調査（厚生労働省）

調査データのトピック別割合



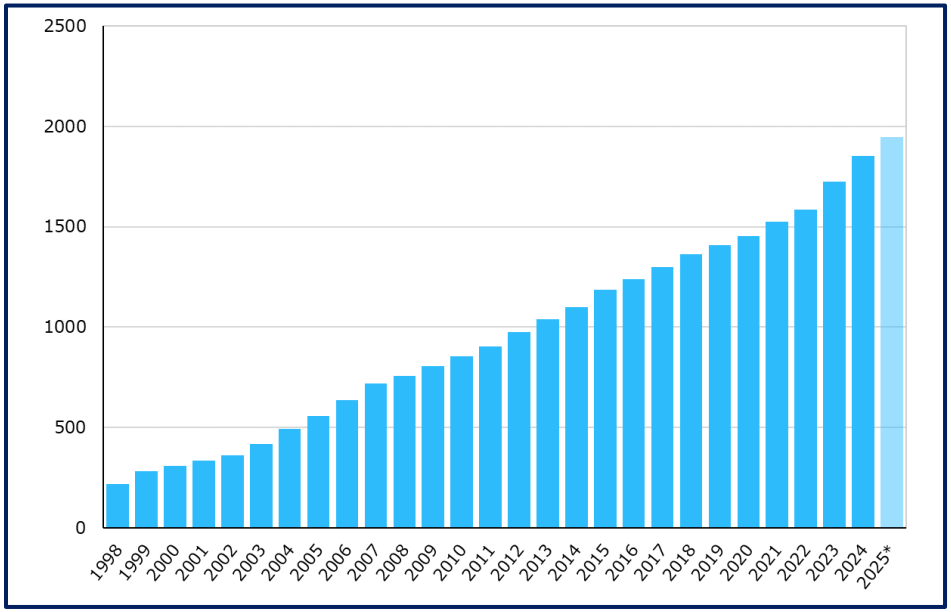
（2026年1月時点）

データ寄託・公開の促進

- データ寄託の促進
  - 個別相談会の実施
  - 寄託者への表彰事業
- データ公開の促進
  - AIを活用したメタデータ作成

寄託者表彰

公開調査データ数の推移

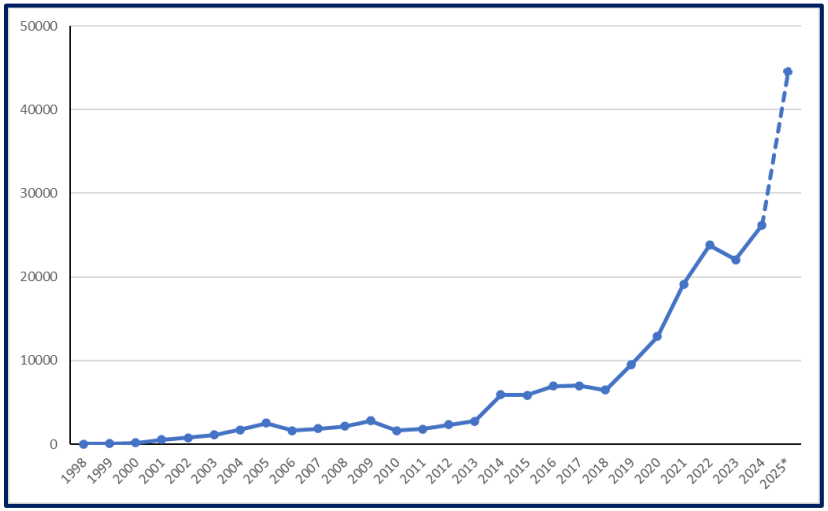


*2026年1月時点

データ利活用の促進

- データの利用方法
  - 利用申請
  - オンライン集計システム (SSJDA Data Analysis)
  - 非制限公開疑似データ

提供データセット総数の推移



*2026年1月時点
（4～12月実績）

オンライン集計システム

SSJDA Data Analysis

東大社研・若年パネル調査(JLPS-Y)wave1-6, 2007-2012(継続調査)

メタデータ

保存済みの出力のみ表示

クロス集計 sex: 問1(1)_性別, ZQ03: 問3_就業状況

N 行% 列%

			問3_就業状況		
			している	していない	合計
問1(1)_性別	male	N	1514	179	1693
		行%	89.4%	10.6%	100.0%
	female	N	1248	426	1674
		行%	74.6%	25.4%	100.0%
	合計	N	2762	605	3367
		行%	82.0%	18.0%	100.0%

	値	自由度	漸近有意確率	正確有意確率
Pearson	126.3558	1	0.0000***	
連続修正	125.3486	1	0.0000***	
尤度比	129.3983	1	0.0000***	
Fisher				0.0000***
ファイ係数	0.1937			
クラメールのV	0.1937			

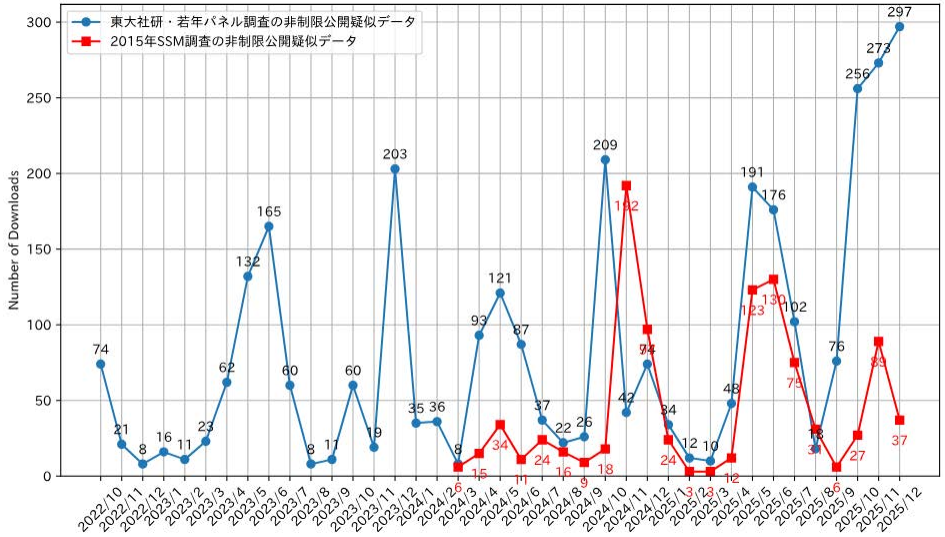
データ利活用の促進

- データの利用方法
  - 利用申請
  - オンライン集計システム (SSJDA Data Analysis)
  - 非制限公開疑似データ

非制限公開疑似データ

名前	ラベル
caseid	
sex	問1(1)_性別
ybirth	問1(2)_生年
mbirth	問1(2)_生月
ZQ03	問3_就業状況
JC_1	問4A(1)_現職・働き方
JC_41	問4A(4-1)_現職・従事時間（1日あたり）

非制限公開疑似データ　利用数の推移



*2026年1月時点

参考：データ分析への生成AIの利用について

利用FAQ

個票データの分析にChatGPTやGoogle Gemini、あるいはその他の生成AIを利用して良いか？ ×

個票データを生成AIに読み込ませて分析を行うことは、生成AIの種類や利用環境（ローカル環境、リモート環境等）に関わらず認められておりません。ただし、個票データと同時に提供される変数ラベルファイルなど、一般公開されている情報を生成AIに読み込ませることは問題ありません。なお、本方針は2025年5月現在のものであり、今後、生成AI技術の動向によって見直していく可能性があります。

https://csrda.iss.u-tokyo.ac.jp/infrastructure/access/access-faq.html

データ利活用の促進

- データ利活用の促進
  - 若手研究者による論文を対象とした表彰事業
  - 二次分析研究の普及および若手研究者の育成のための事業
    - 課題や参加者を公募する「二次分析研究会」
    - 誰でも参加可能な有料の「計量分析セミナー」
  - 社会科学データインフラ・ワークショップを開催

JDCatとの
連携データを利用

二次分析研究会

計量分析セミナー

データインフラ・ワークショップ

JDCat連携の促進

人文学・社会科学総合データカタログ
JDCat Japan Data Catalog for the Humanities and Social Sciences

「人文学・社会科学総合データカタログ（Japan Data Catalog for the Humanities and Social Sciences）」https://jdcat.jsps.go.jp/
連携している人文学・社会科学系のアーカイブ（リポジトリ等）に掲載されている一部の研究データのメタデータを横断的に一括検索することが可能

※2026年6月時点で、およそ16万件のメタデータを公開

トップページ

メタデータの一例（NFRJ18）

家族についての全国調査（第4回全国家族調査，NFRJ18），2019

公開日	2026-04-17
タイトル	
タイトル	家族についての全国調査（第4回全国家族調査，NFRJ18），2019
作成者	日本家族社会学会全国家族調査委員会，実査は(社)中央調査社
配布者	SSJデータアーカイブ
配布者URI	https://csrda.iss.u-tokyo.ac.jp/ssjda/
編集者	SSJデータアーカイブ
所蔵者・寄託者	日本家族社会学会全国家族調査委員会
URI	https://ssjda.iss.u-tokyo.ac.jp/Direct/gaiyo.php?eid=1733
アクセス権	制約付きアクセス
権利情報	https://csrda.iss.u-tokyo.ac.jp/access/condition/
概要	「第4回全国家族調査」(以下，NFRJ18)は，日本家族社会学会全国家族調査委員会が，全国家族調査（National Family Research of Japan）プロジェクト（以下，NFRJプロジェクト）として，1996年度より実施している継続的全国調査の4回目にあたり，2019年1～4月に実施された。 NFRJプロジェクトは，現代日本における家族の動向を正確にとらえることを目的に，(1)全国確率標本による現状把握，(2)継続調査による変動・趨勢分析が可能なデータセットの構築，(3)データの公開・共同利用による共通分析基盤の拡大と他の家族関係統計データの公開促進，という3点を具体的課題としており，1996年度に「第1回全国家族調査」(NFRJ96)を，2003年度に「第2回全国家族調査」(NFRJ03)を，2008年度に「第3回全国家族調査」(NFRJ08)実施し，その結果は複数の調査報告書および学術図書として出版され，それぞれのデータセットは，SSJデータアーカイブを介して公開提供されている（第1回は0191，第2回は0517，第3回は0817）。 今回実施されたNFRJ18は，NFRJ96とNFRJ03とNFRJ08の継続調査として基本的なデザインを踏襲し，日本家族の趨勢を把握することを可能としている。他方，NFRJプロジェクトの一環として改善も加えており，主なものとして「家族形成期の困難」に照準を当てた点が挙げられる。「家族形成期の困難」とは，未婚化，少子化，仕事と家族のバランスの困難，育児負担の問題，若年者の格差と貧困の問題，ひとり親世帯の貧困の問題，高齢者介護，世代間格差など，近年の家族に関連する社会的課題に共通する構同を要約的に捉えた概念である。「家族形成期の困難」の実態を過去の調査よりも更に詳細に把握できるよう，家族形成期のライフコース経験のダイナミクスを捉えるために，パネル調査に代わる工夫として回顧的な調査項目を導入している。
対象時期	2019 - 2019
調査日	2019/01 - 2019/04
母集団	日本国内に居住する1946～1990年生まれの日本国民（2018年12月31日現在で28～72歳）

JDCat連携の促進

人文科学

社会科学

中核機関

東京大学史料編纂所
HISTORIOGRAPHICAL INSTITUTE, The University of Tokyo

東京大学 社会科学研究所

JDCat

奈良文化財研究所
Nara National Research Institute for Cultural Properties

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
Panel Data Research Center at Keio University

神戸大学附属図書館
KOBE UNIVERSITY LIBRARY

一橋大学経済研究所
IER Institute of Economic Research
Hitotsubashi University

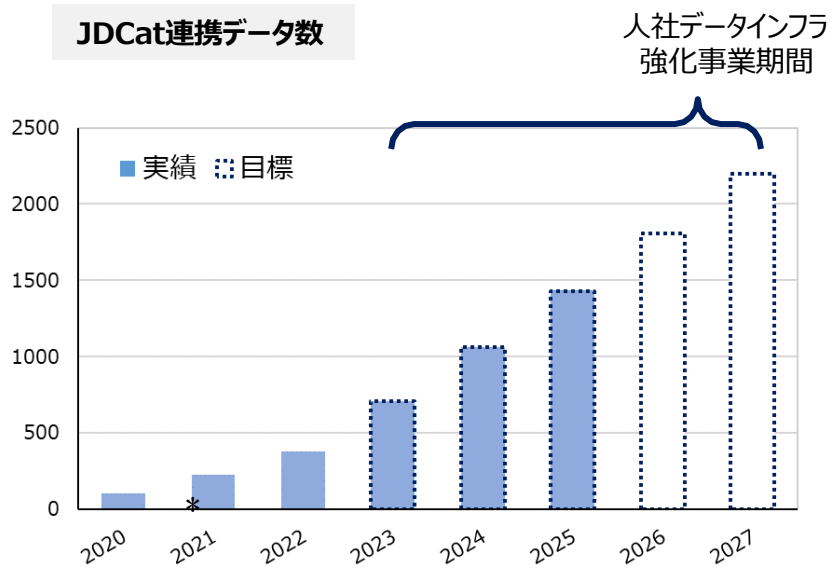
文部科学大臣認定　日本版総合的社会調査共同研究拠点
大阪商業大学 JGSS研究センター

中核機関　HP

Youtube（https://www.hi.u-tokyo.ac.jp/di/）　Instagram（https://www.instagram.com/utokyo_dihs/）

JDCat連携の促進

- 連携の継続と加速
  - メタデータ作成におけるAI活用
- 連携データの利用拡大や共同研究の促進
  - 二次分析研究会や計量分析セミナーなどを開催
- メタデータスキーマや統制語彙に関する最新の動向を調査



*2026年1月時点

学際的・国際的な活動

- CoreTrustSeal (CTS)による認証を取得（2025年6月18日付）
  - データリポジトリの国際的な認証機関
  - 現行の評価体系の下では日本初の認証

CoreTrustSeal Homepage　Certificates Archive

Current CoreTrustSeal certified data repositories

Search Repositories

REPOSITORY	INSTITUTION	WEBSITE	ADDRESS	VALID UNTIL	PDF
Social Science Japan Data Archive (SSJDA)	Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo	https://csrda.iss.u-tokyo.ac.jp/en/ssjda/	7-3-1 Hongo, Bunkyo-ku, Tokyo, JP	2028-06-18	

https://doi.org/10.34894/WN7CKT

メタデータ作成へのAI導入

- 従来のメタデータ作成方法
  - SSJDAのスタッフによる作業
  - 寄託者による作業

➡ **AI活用による省力化・効率化**

メタデータ (SSJDA Direct)

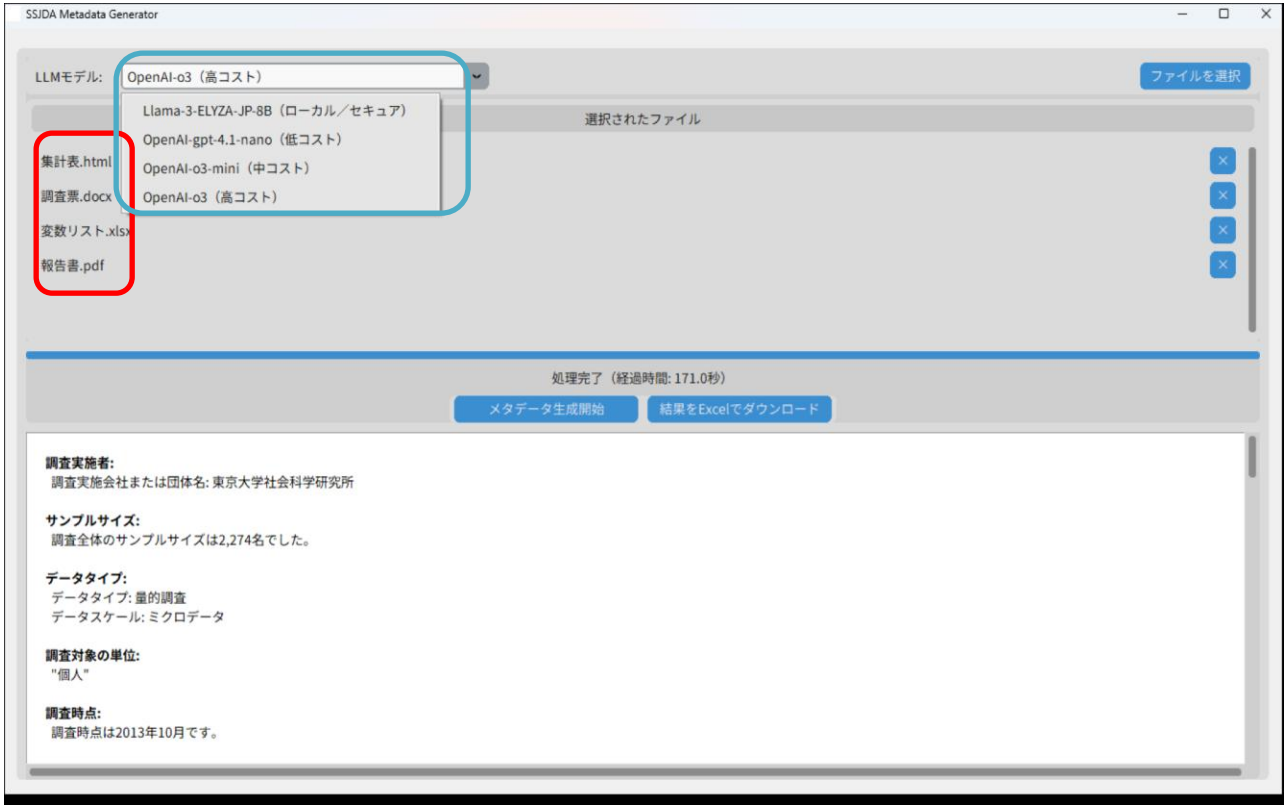
概要	
調査番号	PM050
調査名	東大社研・壮年パネル調査（JLPS-M）Wave 1-5, 2007-2011
寄託者	東京大学社会科学研究所パネル調査プロジェクト
利用申込先・承認手続き	利用方法の詳細はこちら SSJDAが利用申請を承認したときに利用できる
教育目的(授業など)の利用	教育（授業・卒論等）も可
利用期限	研究はなし教育は一年
データ提供方法	ダウンロード
メタデータ閲覧・オンライン分析システムNesstar	利用不可
調査の概要	労働市場の構造変動，急激な少子高齢化，グローバル化の進展などにともない，日本社会における就業，結婚，家族，教育，意識，ライフスタイルのあり方は大きく変化を遂げようとしている。これからの日本社会がどのような方向に進むのかを考える上で，現在生じている変化がどのような原因によるものなのか，あるいはどこが変化してどこが変化していないのかを明確にすることはきわめて重要である。
データタイプ（量的調査／質的調査／官庁統計）	量的調査 量的調査: ミクロデータ
調査対象	日本全国に居住する35～40歳の男女（2006年12月時点）
調査対象の単位	個人
サンプルサイズ	継続調査　回収数：1087　回収率：85% 追加調査　回収数：251　回収率：31%
調査時点	2011-01 ～ 2011-03 2011年1～3月 (wave5)
対象時期	2007 ～ 2011
調査地域	日本 全国
標本抽出	確率: 層別抽出 確率: 多段抽出 層化2段無作為抽出 ・地域（10地域）と都市規模（4類型）の2層により層化（271地点） ・さらに性別・年齢別（5歳階層）に層化 ・抽出台帳は，住民基本台帳を基本とし，住民基本台帳の閲覧が不許可になった地点では，選挙人名簿を使用
調査方法	自記式調査票：紙 調査票による郵送配布，訪問回収法 ・対象者には追跡調査であることを事前に伝えた上で調査に協力を要請し，郵送で調査票を配布。その後記入された調査票を，調査員が訪問し回収。
調査実施者	東京大学社会科学研究所パネル調査プロジェクト 実査は，(社)中央調査社
DOI	10.34500/SSJDA.PM050
寄託者(経費)	（「調査の概要」を参照）

JDCat連携メタデータ

Item type	Harvesting DDI(1)
公開日	2024-01-29
タイトル	
タイトル	東大社研・壮年パネル調査（JLPS-M）Wave 1-5, 2007-2011 Japanese Life Course Panel Survey of the Middle-aged (JLPS-M) Wave 1-5, 2007-2011
作成者	東京大学社会科学研究所パネル調査プロジェクト 実査は，(社)中央調査社
配布者	SSJ データアーカイブ
配布者URI	https://csrda.iss.u-tokyo.ac.jp/ssjda/
配布者	SSJDA
配布者URI	https://csrda.iss.u-tokyo.ac.jp/en/ssjda/
編集者	SSJデータアーカイブ SSJDA
所蔵者・寄託者	東京大学社会科学研究所パネル調査プロジェクト Japanese Life Course Panel Surveys (JLPS) project, Institute of Social Science, The University of Tokyo
URI	https://ssjda.iss.u-tokyo.ac.jp/Direct/gaiyo.php?eid=PM050
URI	https://ssjda.iss.u-tokyo.ac.jp/Direct/gaiyo.php?lang=eng&eid=PM050
アクセス権	制約付きアクセス restricted access
権利情報	https://csrda.iss.u-tokyo.ac.jp/access/condition/
権利情報	https://csrda.iss.u-tokyo.ac.jp/en/access/condition/
概要	労働市場の構造変動，急激な少子高齢化，グローバル化の進展などにともない，日本社会における就業，結婚，家族，教育，意識，ライフスタイルのあり方は大きく変化を遂げようとしている。これからの日本社会がどのような方向に進むのかを考える上で，現在生じている変化がどのような原因によるものなのか，あるいはどこが変化してどこが変化していないのかを明確にすることはきわめて重要である。
対象時期	2007 - 2011
調査日	2011/01 - 2011/03
母集団	日本全国に居住する35～40歳の男女（2006年12月時点） Male and female residents in Japan aged between 35 and 40 in December 2006
観察単位	個人 Individual
サンプリング方法	確率: 層別抽出, 確率: 多段抽出 Probability: Stratified, Probability: Multistage
調査方法	自記式調査票：紙 Self-administered questionnaire: Paper

メタデータ作成へのAI導入

- 開発中のアプリケーション（イメージ）



既存の調査票などの各種ファイル
↓
LLMモデルを選択
↓
メタデータの生成
↓
結果を出力

まとめ：現状と課題

- **1. データ寄託・公開の促進**
  - 寄託を継続・拡大させるための各種取り組み
  - 個票データのキュレーションやメタデータ作成の効率化
- **2. データ利活用の促進**
  - 表彰事業，データの二次的利用に関する教育活動
  - 社会科学データインフラ・ワークショップ、連続セミナーの開講
- **3. JDCat連携の促進**
  - 連携の継続と加速
- **4. 学際的・国際的な活動**
  - 国内外の様々な分野における情報発信・共有
- **5. メタデータ作成へのAI導入**
  - 今後の実装に向けた整備や試験的な運用

データの寄託をお待ちしています
ssjda@iss.u-Tokyo.ac.jp