

社会科学分野におけるデータ提供の実態 ーデータアーカイブ機関と利用者の最前線ー

仲 修平

東京大学社会科学研究所 助教

附属社会調査・データアーカイブ研究センター

2020.10.2

第1回SPARC Japanセミナー2020

本報告の目的

- ・ 目的

次の3点を通じて、社会科学分野におけるデータの公開と制限の境界線を共有する。

- (1) 社会調査・データアーカイブ機関の役割
- (2) 個票データを提供するフローの実態
- (3) 利用者の申請に対する検討の内実

1. 所属機関の役割

2. 保有データ

3. データ提供フロー

4. 提供／制限の内実

5. 制限のケース

6. まとめ

社会科学データの公開に関わる機関

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
Center for Social Research and Data Archives
Institute of Social Science, The University of Tokyo



東大社研パネル調査の調査票（例）

調査内容：働き方とライフスタイルに関する調査
 調査期間：2007年～継続中

問4. あなたの現在および最初に就いたお仕事についてうかがいます（現在働いていない方は、直近の（最後に就いた）お仕事についてうかがいます）。

<現在(直近)のお仕事>と<学校を卒業後、最初に就いたお仕事>が同じ場合は、Aにのみ回答してください。	A.現在(直近)のお仕事	B.学校を卒業後最初に就いたお仕事
(1) 働き方 もっとも近いものを1つ選んでください。(○は1つ)		
1. 経営者、役員	1	1
2. 正社員・正職員	2	2
3. パート・アルバイト(学生アルバイトを含む)・契約・臨時・嘱託	3	3
4. 派遣社員	4	4
5. 請負社員	5	5
6. 自営業主、自由業者	6	6
7. 家族従業者	7	7
8. 内職	8	8
9. その他	9	9
	[9. その他] 具体的に	
(2) お仕事の内容 もっとも近いものを1つ選んでください。(○は1つ)		
1. 専門職・技術職 …… (医師、看護師、弁護士、教師、技術者、デザイナーなど専門的知識・技術を要するもの)	1	1
2. 管理職 …… (企業・官公庁における課長職以上、議員、経営者など)	2	2
3. 事務職 …… (企業・官公庁における一般事務、経理、内勤の営業など)	3	3
4. 販売職 …… (小売・卸売店主、店員、不動産売買、保険外交、外回りのセールスなど)	4	4
5. サービス職 …… (理・美容師、料理人、ウェイトレス、ホームヘルパーなど)	5	5
6. 生産現場職・技能職 …… (製品製造・組立、自動車整備、建設作業員、大工、電気工事、農水産物加工など)	6	6
7. 運輸・保安職 …… (トラック・タクシー運転手、船員、郵便配達、通信士、警察官、消防官、自衛官、警備員など)	7	7
8. その他	8	8

調査地域：全国

調査対象：20-34歳・男女

調査方法：郵送配布

訪問回収

アタック数：約1万人

回収票数：3,367人

<ポイント>

調査対象者の自宅へ訪問！

→1票を入手することに
 労力がかなり掛かっている。

社会科学分野で主に扱う個票データ (例)

個票データ

調査票によって得られた情報を
1行に調査対象1つの情報となるように置き換えたもの

調査対象者の番号

性別

出生年

出生月

	PanelID	CN	sex	ybirth	mbirth
1	1	1	男性 1	1981	2
2	2	1	1	1975	1
3	3	1	1	1967	10
4	4	1	女性 2	1985	5
5	5	1	2	1985	10
6	6	1	2	1977	10
7	7	1	2	1978	5
8	8	1	2	1976	9
9	9	1	2	1974	6

表の見方) 1行目: 男性, 1981年2月生まれ

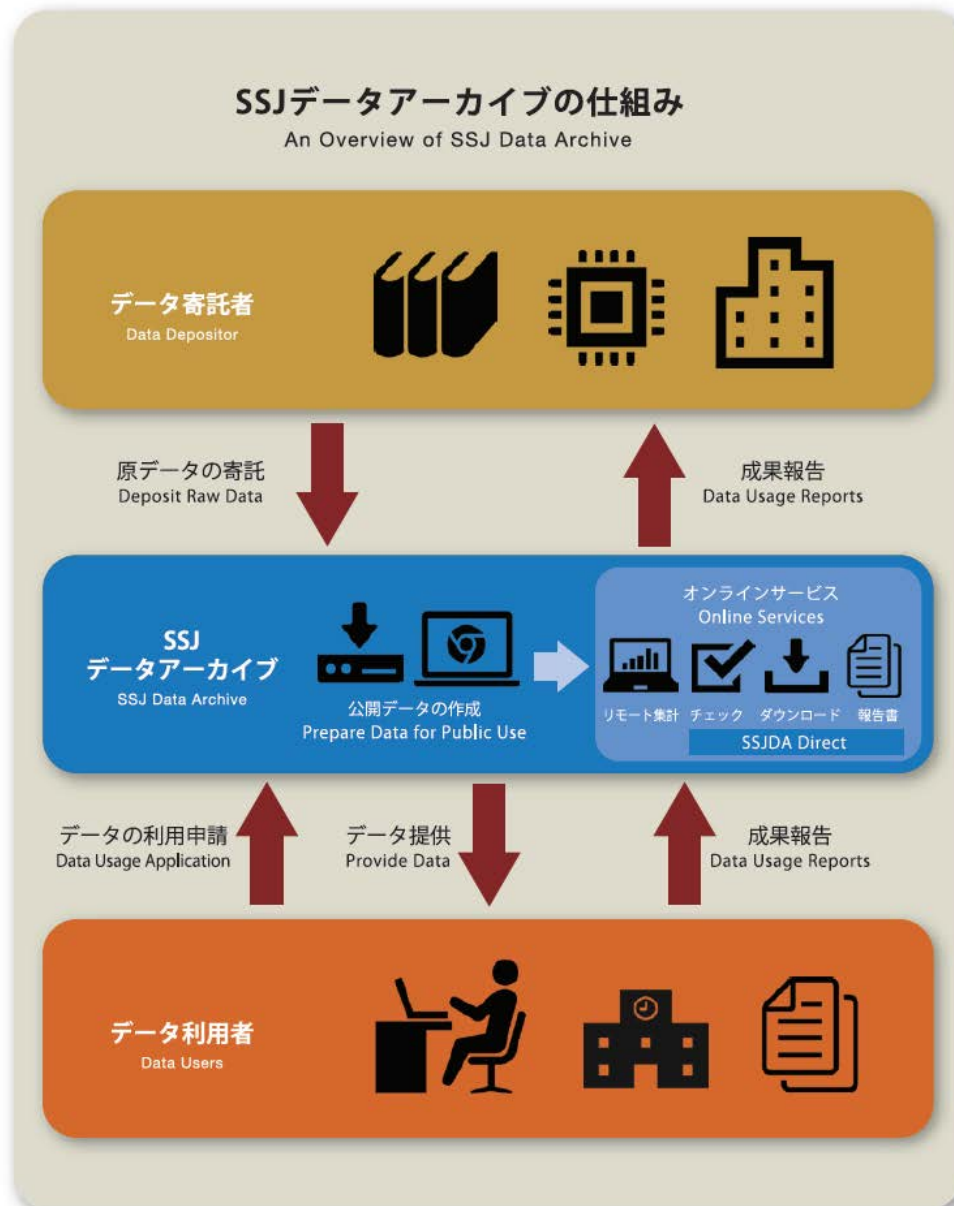
SSJデータアーカイブの仕組み

Social Science Japan
Data Archive

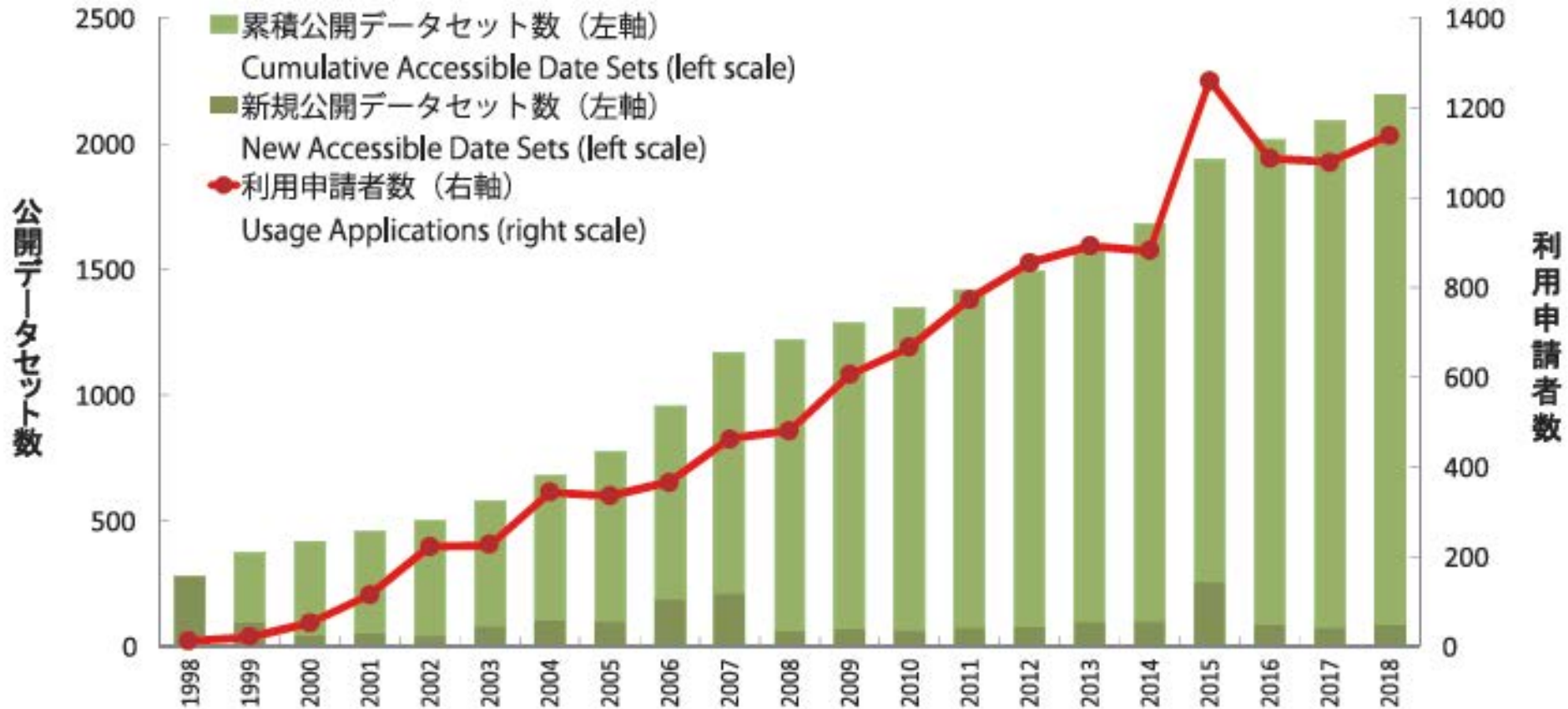
1.各主体が調査を実施

2.各主体が実施した調査
データをSSJDAへ寄託

3.利用者が公開データ
を利用=二次分析



公開データセット数と利用申請者数



新規公開データセット数：70件程度で推移（直近3年平均）

利用申請者数：海外からの申請者も徐々に拡大

SSJDAのデータを利用した研究業績

	2016年	2017年	2018年
著書	12	19	15
論文	41	52	53
学位論文	124	165	161
合計	177	236	229

注：成果物は数年経ってから届く場合があるため、公表値とは異なる可能性がある。

二次分析による研究成果は増加傾向
若手研究者による**学術書の出版**
卒論・修論・博論で活用実績

1.所属機関の役割

2.保有データ

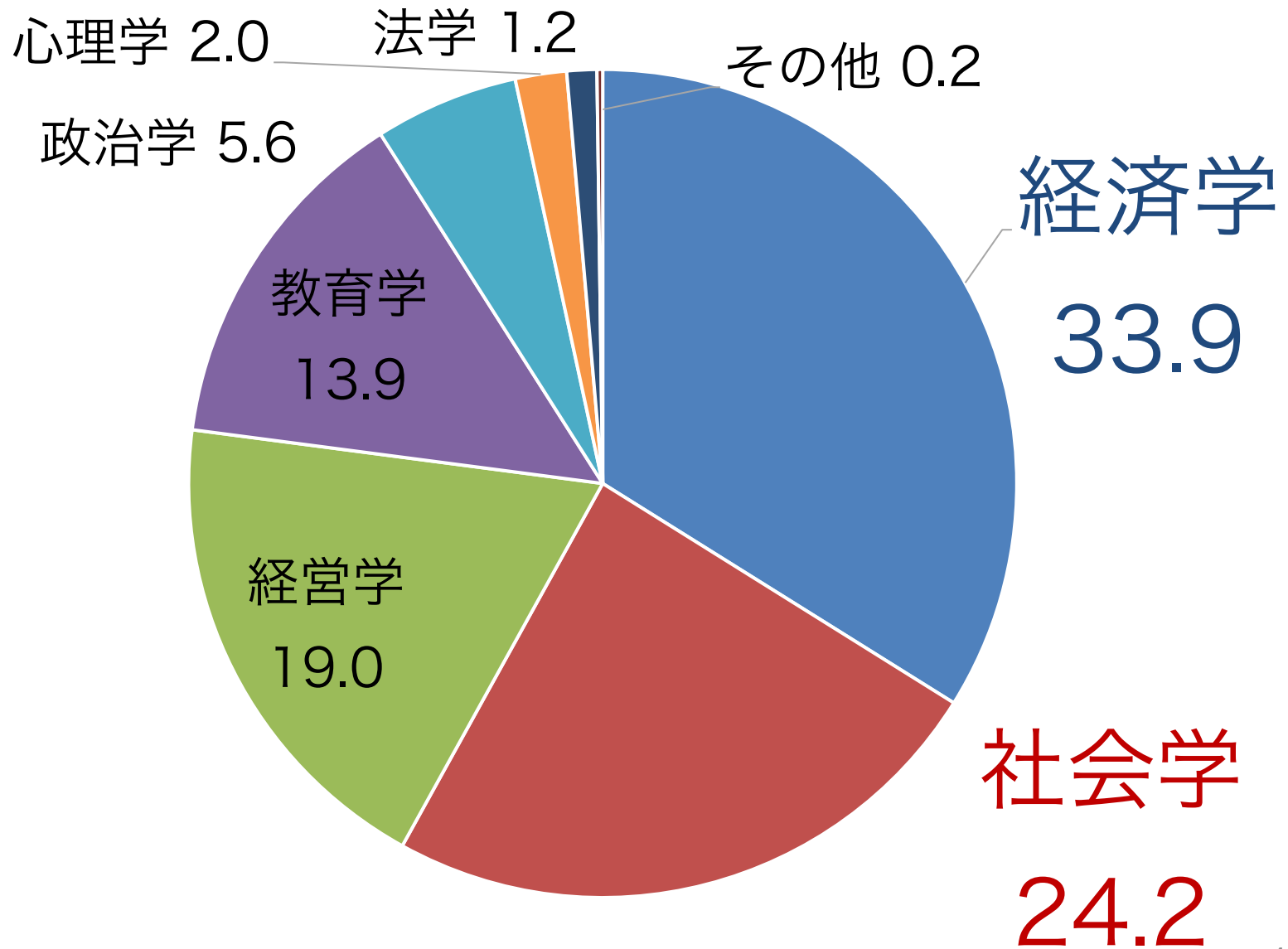
3.データ提供フロー

4.提供／制限の内実

5.制限のケース

6.まとめ

保有データの学問分野



公開・非公開の状況

- ・ 個票データの状況

1. 公開中

2. 公開準備中（データクリーニング等）

3. 保存目的（寄託段階より保存のみ等）

4. その他

保有している多くの個票データは、
公開することを前提としている。

1.所属機関の役割

2.保有データ

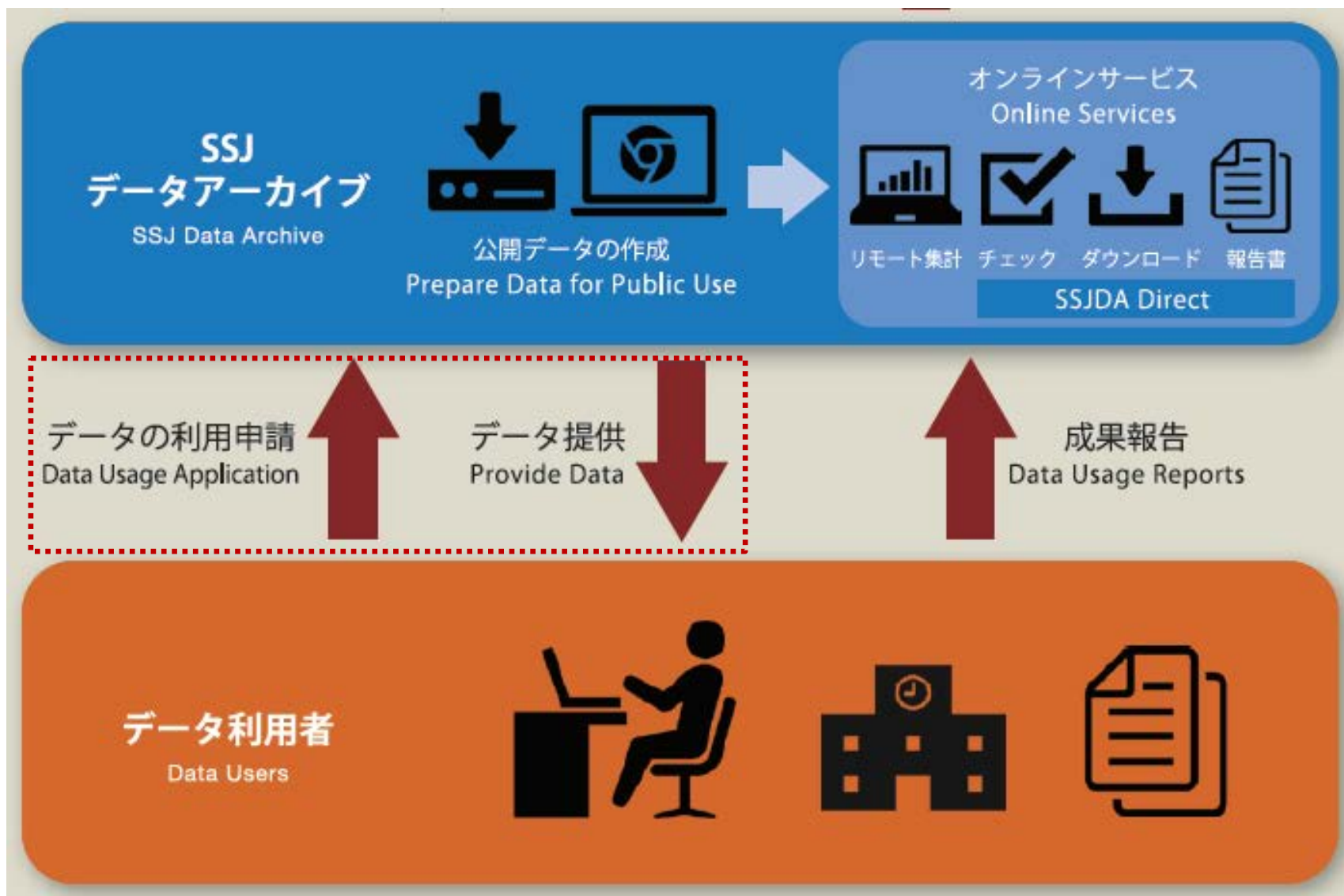
3.データ提供フロー

4.提供／制限の内実

5.制限のケース

6.まとめ

データの提供フロー



データ提供フローの3ステップ

1.利用者による利用申請書の提出

2.SSJDAによる承認の判断

3.データの提供

<申請画面>

利用目的

研究 教育

すべての研究計画を目視によって確認したうえで、承認可否を判断

データ一覧

調査番号	調査名	寄託者名	データ提供方法	利用期限	Nesstar	
PY010	東大社研・若年パネル調査 (JLPS-Y) Wave1 基本データ, 2007	東京大学社会科学研究所パネル調査プロジェクト	ダウンロード	* 期限なし	--	<input type="button" value="選択解除"/>

研究計画 (200文字以上 400文字以内)

[過去の研究計画をコピーする](#)

- 1.所属機関の役割
- 2.保有データ
- 3.データ提供フロー

4.提供／制限の内実

- 5.制限のケース
- 6.まとめ

提供／制限に対する基本スタンス

- ・ 原則は提供

利用目的に合致した場合、
データを利用者へ提供する。

- ・ 制限は例外

制限は**利用条件に満たない場合**を想定

利用承認の判断ポイント

1. **利用条件**に合致しているか？
2. **申請内容**が妥当であるか？
3. **利用状況**が適切であるか？

申請の承認判断①利用条件に合致？

利用資格：大学又は公的研究機関の**研究者**、
教員の指導を受けた**大学院生**

利用状況：過去の申請に対する対応

〈NGとなる場合〉

1.申請者の所属（研究機関・職位等）

民間企業・学部学生・高校教員など

※学部生は指導教員からの申請

2.過去の申請に対する対応状況

利用報告／受講者リストの未提出

申請の承認判断②申請内容の妥当性

- ・ 研究計画の内容

研究／教育の目的は妥当か

【研究】

計画内容が十分に示されているか

計画内容と変数のミスマッチはない

【教育】

卒論・授業名などの記載有無

申請の承認判断③利用状況の適切性

- 一度の申請で大量データを申請
- 学生が教員としてアカウントを作成
- **アカウントの不正利用**
 - 複数名で1つのアカウントを共有
 - 教員が自らのアカウントを学生へ貸与
 - 制限を加えたケースを紹介

- 1.所属機関の役割
- 2.保有データ
- 3.データ提供フロー
- 4.提供／制限の内実

5.制限のケース

- 6.まとめ

最も重い制限を加えた例

- ・ **利用者のアカウントを停止**

- 1.不正利用の発覚

教員が学生へアカウント貸与

- 2.改善を促したが、十分な対応を取らない

- 3.機関内の合意をもとにアカウントを停止

→すぐに停止措置をしたわけではない。

検討を何度も重ねた結果としての停止措置

制限を加えた例

- 学部ゼミメンバーでアカウントを共有
 1. 同じゼミに所属する人たちが
1つのアカウントで利用申請を繰り返す。
 2. 先方への改善依頼
SSJDA内のシステム検討
 3. 2016年3月より
学部学生の申請は指導教員による申請

- 1.所属機関の役割
- 2.保有データ
- 3.データ提供フロー
- 4.提供／制限の内実
- 5.制限のケース

6.まとめ

提供と制限の境界線

- データ提供の原則は常にオープン
- 制限の可否は利用条件の合致を個別に確認
- 重い制限は不誠実な利用者に対してのみ

提供フローの課題

- ・ 個々の申請すべてを目視
 - 人員と時間のコストが大きい
 - 増加する申請件数に対する処理の制約
- ・ 対応案
 - 申請フローの(半)自動化
 - 利用者に対する継続的なセミナー
 - 申請ベースから利用者ベースへ？