

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

国立情報学研究所
朝岡 誠

オープンフォーラム2020
2020年6月10日

発表の流れ

- 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とは
 - 事業の目的と内容
 - ✓ 総合的データカタログ
 - ✓ オンライン分析システム
 - ✓ 拠点機関のアーカイブ機能強化
- 次期JAIRO Cloudへの応用
 - ファセット検索機能
 - オンライン分析システムとの連携
 - 制限公開機能

人文学・社会科学データの共有

• 人文学・社会科学データの特徴

- 人々の意識、行動から、人間の営みや社会事象を把握
 - 調査対象者のプライバシー情報の秘匿化が必要
 - 条件を満たした利用者によりのみデータを提供するスタイル

• オープンサイエンスの流れにより、

- 研究データ共有が研究のスタンダードに
 - ジャーナル投稿や研究資金獲得時に推奨
 - 共有されたデータを用いた研究が一般化
- データ共有のインフラ整備が進展



人文学・社会科学データの共有

•日本の現状

- 機関を通じて、研究データの共有が行われているが
 - 「私のデータ」意識が強く、過去の研究データの死蔵、散逸
 - 多言語対応の遅れ、海外研究者との共同研究機会の喪失
 - データ共有環境の整備、およびガイドライン整備の遅れ
- 人社系データインフラの遅れに対する危機感



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(人社データインフラ事業)

- 日本学術振興会による事業
- 2018年度から2022年度の5年間で
 - ✓ 研究データ共有システムを整備
 - ✓ 拠点機関に研究データの整備を委託

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

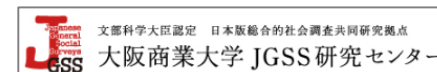
中核

- 研究データインフラの開発・整備
 - 総合的データカタログ
 - オンライン分析システム
- 研究データ公開ガイドラインの整備・周知
→研究データ共有のためのシステム整備



拠点機関

- 研究データの収集、アーカイブ機能強化
- メタデータの作成、多言語化
- 研究データ公開支援
→共有される研究データの整備



総合的データカタログ

2021年4月 一般公開

CiNii Research

検索



いろいろな研究者
が行った社会調査
が検索できる！

総合的
データカタログ



海外のデータ
検索サイト

国際標準メタデータ
DDIによるメタデー
タ自動収集機能



日本の人文学社会科学データの公開サイト

総合的データカタログ

Language: English

wekosoftware@nii.ac.jp

Top

Workflow

Communities

ファセット検索

トピック

☐ 労働と雇用(693)
☐ 人口統計学と人口(248)
☒ 経済(723)
☐ 健康(524)

アクセス制限

☐ 一般公開(744)
☐ 制限公開(1256)

データタイプ

☐ 量的調査(1283)
☐ 政府統計(754)

Type and press enter to search

Full text

Keyword

Search Results

1 - 2 of 2 results.

Export

Display Order: ID asc Display Number 20

日本版General Social Surveys <JGSS-2008>

大阪商業大学JGSS研究センター、大阪商業大学JGSS研究センター、日本版General Social Surveys (JGSS) は、アメリカのNational Opinion Research Centerが1972年から実施しているGeneral Social Survey (GSS) に範を取り、国際比較を視野に入れて、日本社会と人々の意識や行動の実態を把握することを目的としたプロジェクトである。

Show source

東大社研・若年パネル調査 (JLPS-Y) Wave1-2, 2008

東京大学社会科学研究所パネル調査プロジェクト、SSJデータアーカイブ、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ともないう、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。

Show source

トピック、アクセス制限、
データタイプ等で絞り込み可能

検索内容に応じて、各機関の最新情報をもとに研究データが検索できる
→詳細情報から公開サイトにアクセス可

7

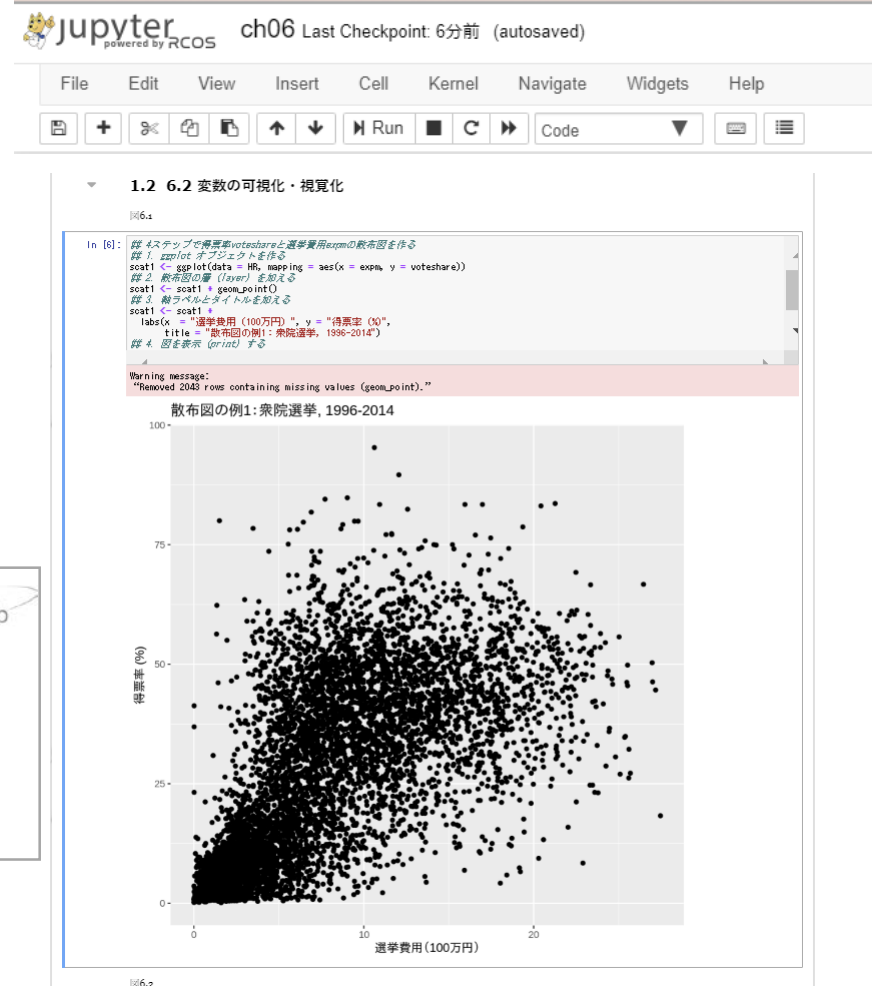
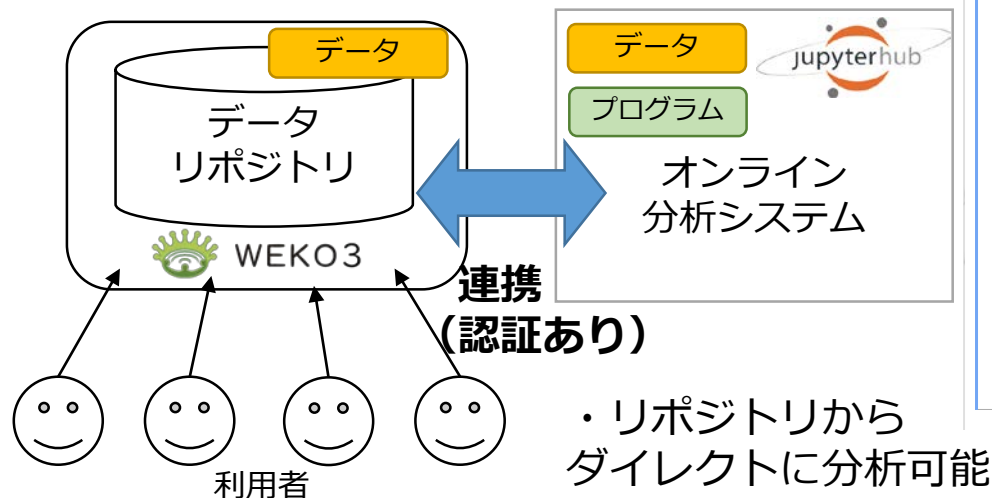
オンライン分析システム

2022年 一般公開

JupyterHub

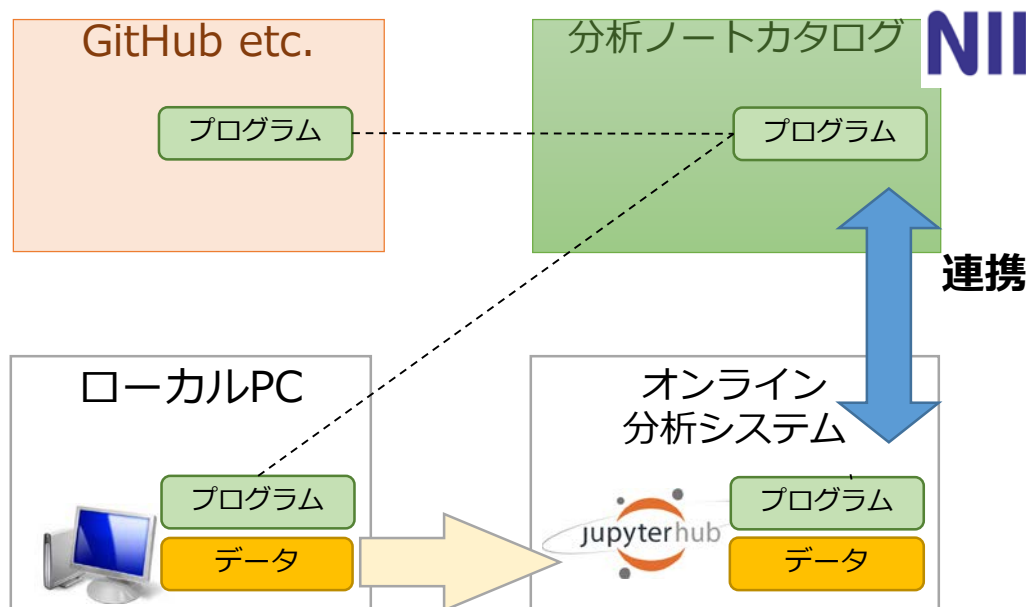
Webブラウザ上から、学認を使ってアクセス

- ・研究データ、分析プログラムをアップロードが可能
- ・R、Pythonを用いたデータ分析が可能



分析ノートカタログ

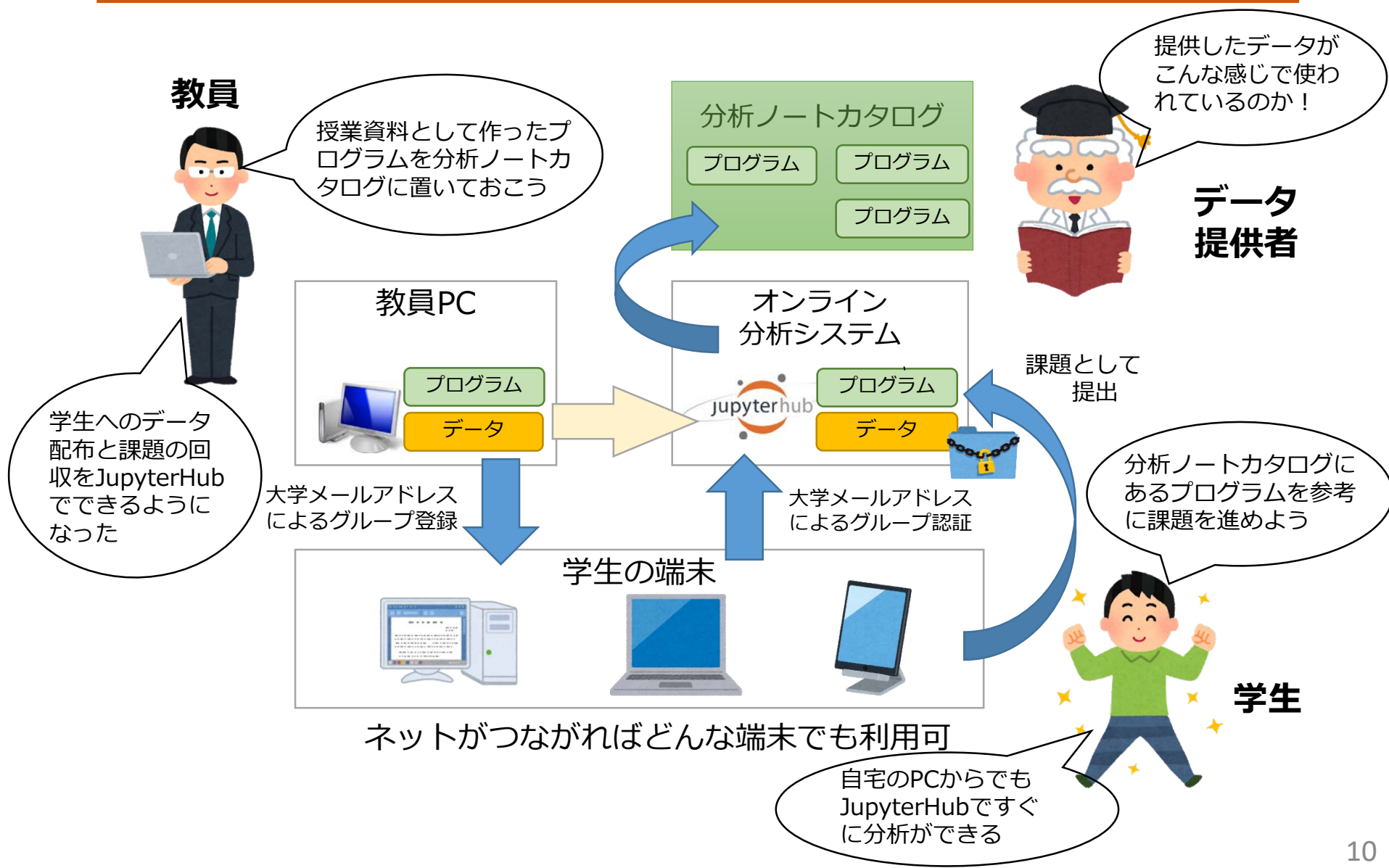
オンライン分析システムで作ったプログラムを集めた「**分析ノートカタログ**」を構築予定
 分析プログラムを検索できる環境を提供する予定



自分の分析を正しく
 解釈してもらえた！

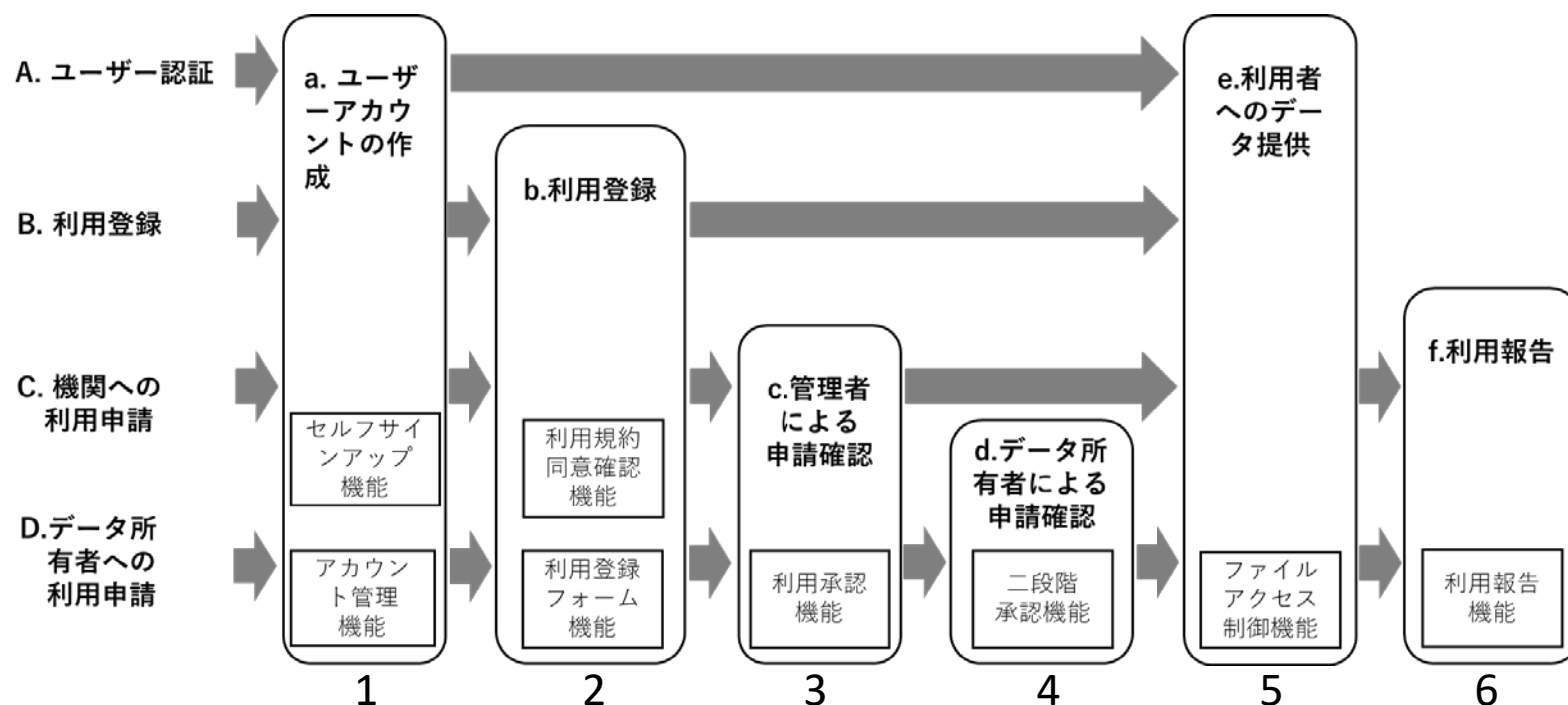
先行研究で使われた
 分析プログラムが見
 つかった！

教育への応用



拠点機関のアーカイブ機能強化

- 社会科学分野の研究データを共有するために
制限公開機能を実装



制限公開

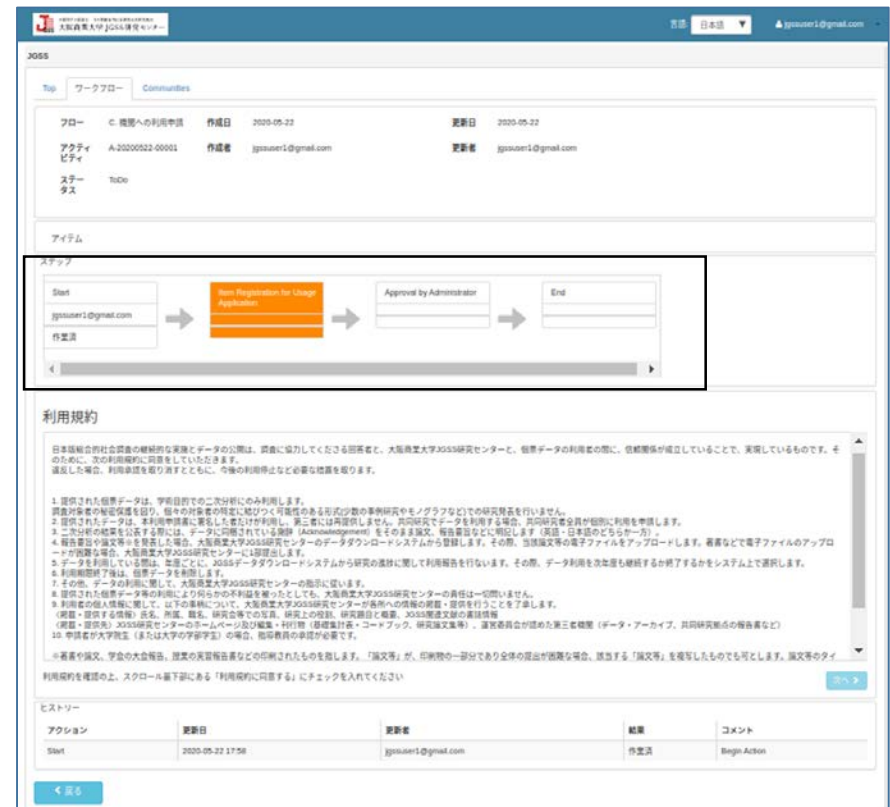
条件を満たした利用者へのみデータを提供する方法。複数のフローから成り立ち、運用にはシステムによるサポートが必要。

制限公開機能

- 拠点機関 大阪商業大学 JGSS研究センターとの共同開発によって、利用申請システムを実装



アイテム画面



利用登録画面

人社データインフラ事業では

人社分野のデータ提供機関の意見をもとにインフラを整備

総合的データカタログ

- ・ 複数のリポジトリデータのポータル検索
- ・ 絞り込み検索（ファセット検索）機能の実装

オンライン分析システム

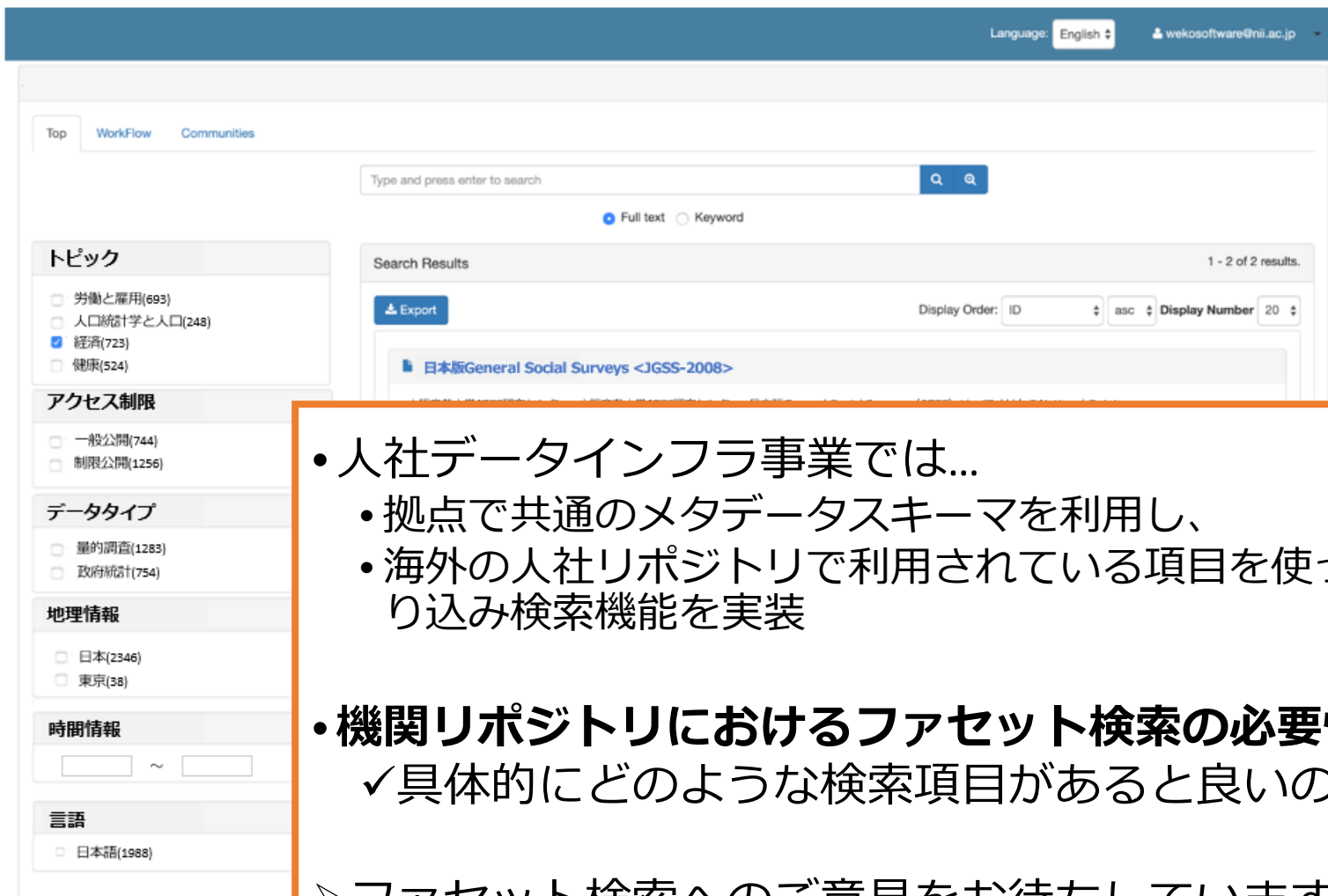
- ・ JupyterHubによる分析システム
- ・ リポジトリ（JAIRO Cloud）との連携

拠点機関のアーカイブ機能強化

- ・ 制限公開機能の実装

次期JAIRO Cloudへの応用①

ファセット検索機能



- 人社データインフラ事業では...
 - 拠点で共通のメタデータスキーマを利用し、
 - 海外の人社リポジトリで利用されている項目を使った絞り込み検索機能を実装
 - **機関リポジトリにおけるファセット検索の必要性**
 - ✓ 具体的にどのような検索項目があると良いのか？
- ファセット検索へのご意見をお待ちしています

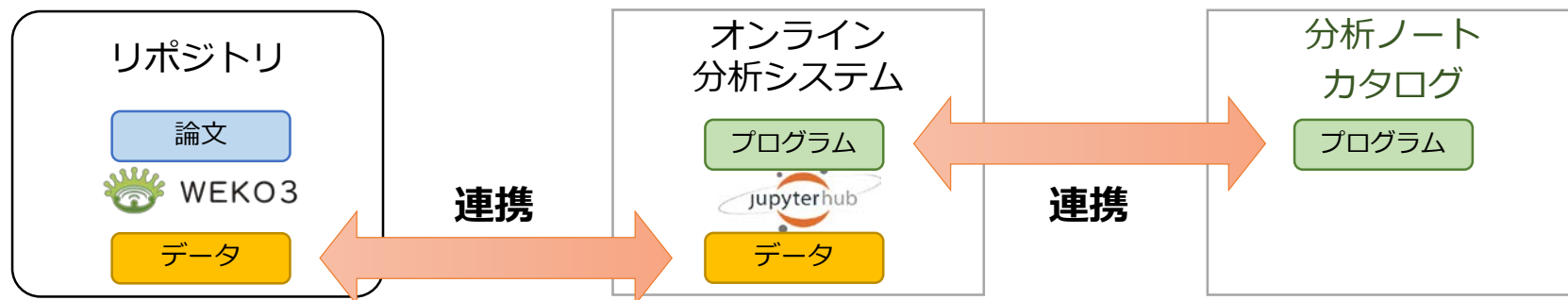
アンケートにご協力をお願いします

https://www.nii.ac.jp/openforum/2020/day3_content2.html

14:40-15:00	研究データ公開に向けたワークフローの検討 北海道大学附属図書館研究支援課	三上 絢子	 講演資料
15:30-15:50	Towards the Remote Curation of CERN INSPIRE 北海道大学附属図書館管理課	前田 隼	 講演資料
15:50-16:10	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 国立情報学研究所オープンサイエンス基盤研究センター	朝岡 誠	 講演資料 <u>アンケートにご協力をお願いします (6/17×)</u>
16:10-16:30	SCPJの移行 神戸大学附属図書館情報管理課	松村 友花	 講演資料
16:30-16:55	質疑応答 全員		

次期JAIRO Cloudへの応用②

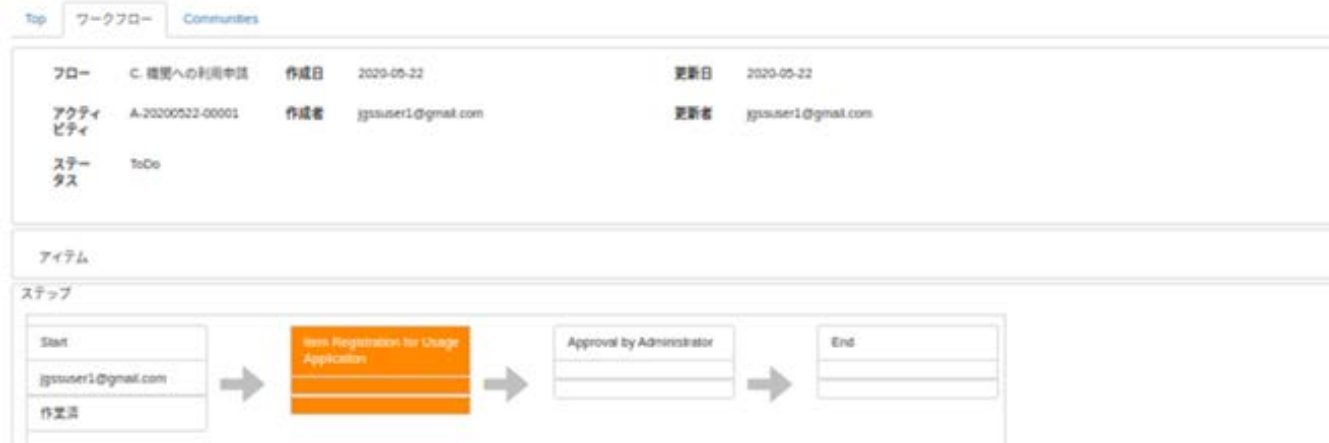
オンライン分析システム連携



- 人社データインフラ事業では...
 - オンライン分析システムを軸に、研究データと分析プログラムを紐付けることで利活用を促進
 - 登録されたコンテンツの教育利用の推進
 - **機関リポジトリとのシステム連携テスト**
 - ✓ 研究データと分析手法を紐付けた公開
 - ✓ オンライン分析システムの授業利用への活用
- 試験運用にご協力いただける機関を募集します

次期JAIRO Cloudへの応用③

制限公開機能



- 人社データインフラ事業では...
 - 拠点機関の公開ニーズへの対応
 - 制限公開機能による申請ダウンロードシステムの構築
 - **機関リポジトリによる制限公開機能の試験運用**
 - ✓ エビデンスデータ等の制限公開データの運用
 - ✓ 制限公開機能を利用したデータ提供
- 試験運用にご協力いただける機関を募集します

まとめ

人社データインフラ事業で開発した以下の機能を次期JAIRO Cloudに組み込むことを計画しています。

- ・ファセット検索機能
- ・オンライン分析システムとの連携
- ・制限公開機能

RCOS

<https://rcos.nii.ac.jp/>
asaoka@nii.ac.jp

これらの機能を試験運用されたい方は
NII RCOSにご連絡ください！

アンケートにご協力をお願いします

https://www.nii.ac.jp/openforum/2020/day3_content2.html

14:40-15:00	研究データ公開に向けたワークフローの検討 北海道大学附属図書館研究支援課	三上 絢子	講演資料
15:30-15:50	Towards the Remote Curation of CERN INSPIRE 北海道大学附属図書館管理課	前田 隼	講演資料
15:50-16:10	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 国立情報学研究所オープンサイエンス基盤研究センター	朝岡 誠	講演資料
16:10-16:30	SCPJの移行 神戸大学附属図書館情報管理課	松村 友花	講演資料
16:30-16:55	質疑応答 全員		

[アンケートにご協力をお願いします \(6/17×\)](#)