

国内社会科学系データアーカイブの横断検索に向けての試行と検討

国立情報学研究所 船守美穂, 林正治, 加藤文彦 東京大学 三輪哲
立教大学 朝岡誠, 高橋かおり 関西学院大学 前田豊

問題の所在

国内複数のSSDAが横断検索できない。



解決の方向性

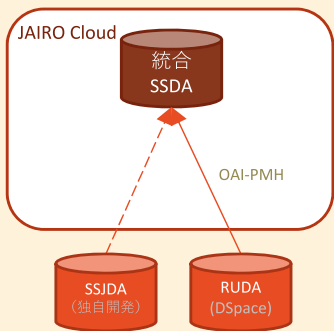
国立情報学研究所(NII)の学術情報流通サービスを利用した、横断型SSDAの実現

- 機関リポジトリクラウドサービス
- NII学術情報ナビゲータ



単一リポジトリへの統合方法とその特性

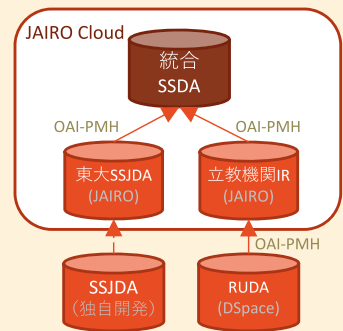
直接登録方式



- : シンプル
- ×: 統合SSDA管理者が必要

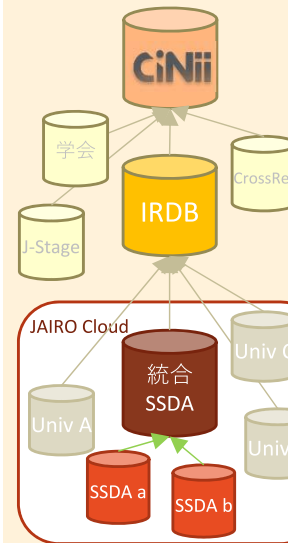


機関リポジトリ経由方式



- : 各SSDAや大学図書館の管理体制を利用可能
- : 第2段階でOAI-PMH利用可能
- ×: 各機関で2リポジトリ管理要
- ×: JAIRO Cloudの追加費用発生

アーキテクチャの検討



各種学術リソース横断型検索 (CiNii)

- : CiNiiブランド、○: 学際領域の利用
- : 論文, データ等リンク情報利用可能
- ×: 現段階ではデータのサポート無
- ×: 社会科学に特化した検索不可

機関リポジトリ横断型検索 (IRDB)

- : OAリソースを検索可能
- ×: 機関IRのみの検索は意味無し
- ×: IRDB自体が知られていない

SSDA特化型検索 (統合SSDA)

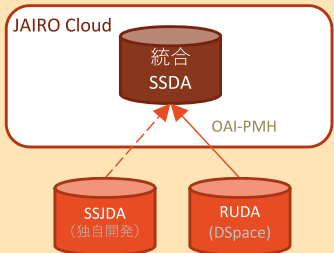
- : 社会科学分野研究者に有用
- : 社会科学詳細検索の開発可能
- ×: 社会科学外では利用が少ない

単一SSDA内検索

- : 当該SSDAのアピールに有用
- ×: SSDA横断型検索が実現不能

パイロットスタディ

- 東京大学SSJDA
- 立教大学RUDA
- 国立情報学研究所 JAIRO Cloud



※ CiNiiが研究データに未対応のため、各SSDAのメタデータレベルの統合までのみでパイロットスタディ実施。

メタデータ (DDI準拠) 全体

File / Name	調査時期
License	観察単位
調査票 (DL可)	調査地域
	調査対象国
アイテムタイプ	調査最小範囲水準
言語	標本抽出法
キーワード	標本サイズ
調査番号	有効回収数
調査略称	有効回収率
調査主体	調査方法
	主要変数
調査資金	関連資料
寄託者	
配布者	
調査概要	データ利用条件
研究分野	引用時の表記
母集団	リンク

変数情報

アイテムタイプ	記述統計
言語	
質問文	回答形式
質問内容	調査タイトル
回答項目	調査番号
	調査主体
	配布者
欠損値指定	

パイロットスタディにより見えてきた課題

統合SSDA管理者の必要性

- ✓ 情報システムにより管理業務を最低限にまで簡素化しても、運用方針の策定や、各SSDAからの質問等への対応などの管理業務は発生する。
- ✓ 統合SSDA管理の仕組みオプション
 - ① 運営委員会方式、② 専属組織の設置

共通のメタデータスキーマの必要性

- ✓ 複数のSSDAを統合する際に共通のメタデータスキーマが必要であるが、現在多くのSSDAは独自規格を使用。かつ、国際標準であるDDIは国内であまり利用されていない。
- ✓ メタデータスキーマのオプション
 - ① DDI準拠のスキーマ
 - ② Dublin Core準拠 (10項目程度)
 - ③ 独自規格

登録内容の標準化に関わる課題

- ✓ 共通のメタデータを利用していても、各SSDAにおける運用により、同じデータ項目でも入力される内容が異なる場合がある。
- ✓ 入力内容を標準化した方が検索精度は上がるが、標準化の労力は膨大。



国レベルの新たな動き

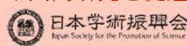
統合イノベーション戦略

日本を「世界で最もイノベーションに適した国」とするための3つの「知の源泉」の一つに、オープンサイエンスを位置づけ。



人文学・社会科学データインフラストラクチャー構築プログラム

データを分野や国を超えて共有・利活用する総合的な基盤を構築することにより、研究者のデータ共有や共同研究を促進。



国の動きと連携しながら進めます!

NII研究データ基盤

学術研究活動の過程で生成される研究データや関連の資料を管理・公開・検索するための情報インフラを提供することを通じて、日本の学術をオープンサイエンスへと導く。

