

資料3

国際的動向を踏まえたオープン
サイエンスの推進に関する検討会
(第8回)
平成30年12月27日(木)

NII Research Data Cloud の開発状況

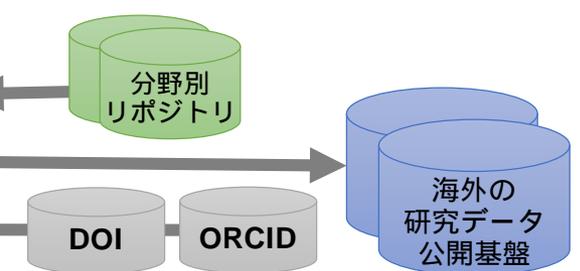
国立情報学研究所
山地一禎

2018年12月27日

国際的動向を踏まえたオープンサイエンスの推進に関する検討会

NII Research Data Cloud

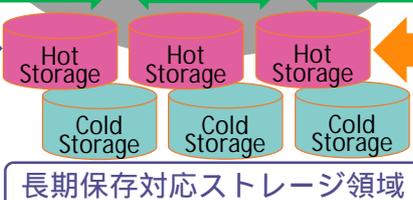
- 機関リポジトリ + 分野別リポジトリやデータリポジトリとも連携
- 研究者や所属機関、研究プロジェクトの情報とも関連付けた知識ベースを形成
- 研究者による発見のプロセスをサポート



2020運用開始

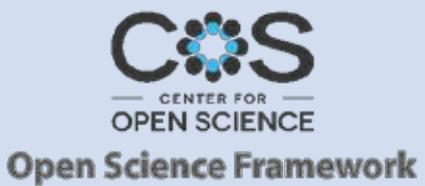


- データ収集装置や解析用計算機とも連携
- 研究遂行中の研究データなどを共同研究者間やラボ内で共有・管理
- 組織が提供するストレージに接続した利用が可能



- データ管理基盤における簡便な操作で研究成果の公開が可能
- 図書館員やデータキュレータによる、メタデータや公開レベル統計情報などの管理機能の提供

国際協力体制

	データ公開基盤	データ検索基盤	データ管理基盤
海外連携先		 OpenAIRE など	 Open Science Framework
連携内容	<ul style="list-style-type: none"> • Invenioは、CERNが提供する複数のデータベースの基礎となるシステム • Invenioをベースに公開基盤を構築 • ソースコードの提供 • 内部ドキュメントの共有 • 定期的なTV会議によるプロジェクト連携 	<ul style="list-style-type: none"> • OpenAIREは、H2020のサポートで運用されるディスカバリーサービス • メタデータ連携 • テキストマイニングに関する技術的連携 • ナレッジグラフ構築に関する技術的連携 	<ul style="list-style-type: none"> • Open Science Framework (OSF)は、米国NPOのCOSが開発・提供するサービス • OSFを機能拡張し管理基盤を構築 • ソースコードの提供 • 内部ドキュメントの共有 • プロジェクト管理システム上での連携

公開基盤の開発状況

論文リポジトリとしての機能継承

- 総機能数：全242（51機能グループ）2018年度未完了予定
- オリジナルInvenioとの比較：

	モジュール数	コード行数
Invenio	26	61,448
WEKO	25	45,242

データリポジトリとしての展開

- Invenio3：ZENODOとしての集中型データリポジトリを対象
- WEKO3：クラウド型データリポジトリとして機能提供



異なる粒度でのデータリポジトリが必要とされる

持続性

マルチテナント機能

機関リポジトリ内に性質の異なるリポジトリ（テナント）を運用可能

クラウド型データリポジトリの重要機能

柔軟性

ワークフロー機能

コンテンツ種別や運用パターンごとに異なる複数の登録フローを定義可能

公開基盤の外部連携：CERN

CERN Accelerating science
Sign In | Directory


ABOUT
NEWS
SCIENCE
RESOURCES
SEARCH | EN -

News · News · Topic: Knowledge sharing

[Voir en français](#)

CERN collaborates with Japan's NII on digital libraries

15 OCTOBER, 2018 | By Abha Eli Phoboo

How we produce, share and preserve knowledge has transformed with advancements in technology and changes in publishing methods, especially in research. Particle physicists rely heavily on digital library platforms to archive, share and preserve research output. These platforms depend on [Invenio](#), a digital library framework CERN has been developing since 2002.

This year, on 4 September, CERN signed a Memorandum of Understanding with Japan's [National Institute of Informatics \(NII\)](#). The event signals the beginning of a collaboration between the two organisations to extend Invenio with [Next Generation Repository](#) capabilities as defined by the [Confederation of Open Access Repositories](#).

"The collaboration focuses on extending Invenio to become fully Next Generation compliant, which will consolidate CERN technology for digital repositories as the most modern repository framework in the market, making it more useful for any institution worldwide with the same needs. To have Invenio technology developed by two such important institutions will enable an efficient and diverse evolution of the software. This also highlights CERN's capacity to generate technological value for society," says Jose Benito Gonzalez Lopez, project coordinator.

Also on Knowledge sharing



Making CERN's exhibitions more accessible

Knowledge sharing | News | 18 December, 18

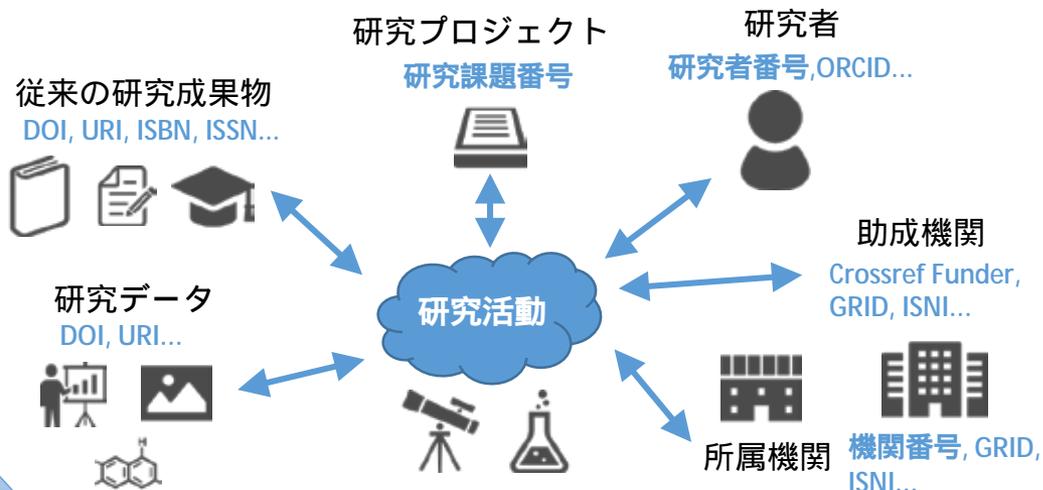


CESP inspires young entrepreneurs to dream big

Knowledge sharing | News | 6 December, 18

検索基盤の開発状況

CiNii Knowledge Base



種別	件数	情報源
学術論文	37,376,419	CiNii Articles, J-STAGE, 機関リポジトリ...
書籍	11,801,960	NACSIS-CAT
博士論文	634,467	NDLオンライン, 機関リポジトリ...
研究プロジェクト	848,051	KAKEN
研究データ	12,568,782	DataCite, Japan Link Center
研究者	2,874,337	KAKEN, NACSIS-CAT
計	66,104,016	
関係リンク	10,813,948	CiNii Articles, KAKEN

検索基盤の外部連携



国際協調によるデータ整備

- 検索基盤では国内の研究データ・関係リンクを整備
- 各国・各地域の主要リポジトリ・ディスカバリーとのデータ交換によってグローバルな発見を実現する
- 異なるメタデータフォーマット間の相互運用性を担保

管理基盤の開発状況


GakuNin RDM 実証実験における
 オープンサイエンス・研究公正・研究活動サポート・システム改良
 の各側面における開発ロードマップ

目的別	2018 Q3	2018 Q4	2019 Q1	2019 Q2	2019 Q3	2019 Q4	2020 Q1	2020 Q2	2020 Q3
オープンサイエンス				解析基盤連携		公開基盤連携		検索基盤連携	四基盤統合
研究公正	研究証跡保全			不正画像検出		不正文書検出		電子実験ノート連携?	
研究活動サポート		ワークフロー機能	グループ機能	データ解析機能	UI改修	デスクトップクライアント	実験装置連携?	HPC連携?	
システム改良	機関管理者向け機能		多要素認証		日本語対応		機関テンプレート機能		

(グレーで「?」の箇所はパートナーが必要なため構想段階)

3ヶ月毎にシステムを拡張
 新機能を即座に全国的に提供

管理基盤の外部連携：北大

学際大規模計算機システム (北海道大学ハイパフォーマンスインタークラウド)



システム利用予定の研究事例

人獣共通感染症リサーチセンター

新興感染症の出現は人類の脅威となっています。人、野生動物、家畜、飼育動物の種体に含まれる遺伝子の大量な塩基配列デー

化学反応創成研究拠点 (WPI-ICReDD)

計算科学・情報科学・実験科学の三分野融合により、人類が未来を生き抜く上で必要不可欠な化学反応を、新たに創成したネットワーク

SINET5と国立情報学研究所 (NII)

学際大規模計算機システムでは、国立情報学研究所 (NII) が提供するSINET5との連携を強く意識して設計しています。特にインタークラウドシステムに関しては、学内ファイアウォール装置を迂回する100Gbpsのバイパス線を設置しており、SINET5が提供するL2VPNおよびL2DDサービスを經由して他拠点とインタークラウドシステムを接続する場合には、SINET5が持つ超広帯域、低遅延などの特長を活用できます。本機能は、東京大学・大阪大学・九州大学などの遠隔サイトへの接続に実際に活用しており、他大学・他の学術機関との接続にも活用することができます。

また、国立情報学研究所が主催する下記のプロジェクトへの参加

研究データ管理基盤



近年、研究論文の成果としての公開だけではなく、研究のもとになったデータやソフトウェアも公開するオープンサイエンス化の流れが急速に進行しています。管理基盤 (GakuNin RDM) は、研究プロジェクト実施中に、個人の研究者あるいは研究グループが研究データや関連資料を管理するための基盤です。

本学の学際大規模計算機システムとの連携では、クラウドストレージをGakuNin RDMのエクストラストレージとして提供する予定です。

また、HPCを利用したデータ



北大クラウド (ストレージ) の活用事例として管理基盤を紹介

管理基盤の外部連携：名古屋大

まとめ

- 名古屋大学における研究データ管理に関するこれまでの取り組み



国立大学経営改革促進事業「研究データの大学間相互利用に向けたアカデミッククラウドの構築」の一環としてGakuNin RDMを活用

管理基盤の実証実験

• 目的

- GakuNin RDMの機能や性能の実運用性の確認
- システム導入担当者様と協力した環境の整備
- ロングランテスト（約1年半利用）
 - 現実的な研究データ管理の課題を洗いだし
 - サービス向上のための開発および運用改善にフィードバック

• 方法

- 1人あたり100GBのNIIストレージをデフォルトで提供
- 参加機関は学内サービスのクラウドストレージをエクストラストレージとして接続可能
- NII学術情報基盤ミーティング2018（全国6か所で開催）にて参加者募集

• 期間

- 2018年 12月 アナウンス開始
- 2018年1～3月 先行試用を開始
 - 協力機関による負荷試験
- 2019年4月～2020年9月
 - ロングラン実証実験

実証実験ML参加機関数
70機関！

海外でのNII RDCの紹介状況

- Yamaji, K.: “NII Research Data Cloud”, Libsense Repository Workshop, Zanzibar, 2018-11.
- Yamaji, K.: “Development on Research Data Infrastructure in Japan”, Panel Discussion during Queen’s University’s Principal’s Symposium, Kingston, 2018-11
- Kato, F., Kanazawa, T., Kurakawa, K., Ohmukai, I.: “Development of Japanese Research Data Discovery Service”, International Workshop on Data Science 2018, Mishima, 2018-11.
- Yamaji, K.: “ Developing E-Infrastructure for Open Science in Japan”, Panel Discussion during Monash Research Data Day, Kuala Lumpur, 2018-10
- Komiyama, Y., Yamaji, K.: “Interdisciplinary research data management service for the whole universities and research institutions in Japan that emphasizes research integrity”, Digital Infrastructures for Research 2018, Lisbon, 2018-10.
- Kato, F., Ohmukai, I., Kanazawa, T., Kurakawa, K.: “CiNii Research: A prototype of Japanese Research Data Discovery”, eResearch Australasia 2018, Melbourne, 2018-10.
- Hayashi, M., Yamaji, K., Hayashi, Y., Tanabe, K., Aoyama, T., Namiki, T., Ikeda, D.: “Historical Transition of Repository Functions and Latest Trends of Repository System Architecture”, 13th International Conference on Open Repositories, Bozeman, 2018-6.
- Yamaji, K.: “Best Practices in Developing and Enhancing Institutional Repositories”, Workshop on Next Generation Repositories in Malaysia, Perlis, 2018-4
- Komiyama, Y., Yamaji, K.: “Nationwide Research Data Management Service of Japan in the Open Science Era”, th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI) , vol.2017, no., pp.129-133, 2018-2.

海外でのNII RDCの紹介状況

- Yamaji, K.: “NII Research Data Cloud”, Libsense Repository Workshop, Zanzibar, 2018-11.
- Yamaji, K.: “Development on Research Data Infrastructure in Japan”, Panel Discussion during Queen’s University’s Principal’s Symposium, Kingston, 2018-11
- Kato, F., Kanazawa, T., Kurakawa, K., Ohmukai, I.: “Development of Japanese Research Data Discovery Service”, International Workshop on Data Science 2018, Mishima, 2018-11.
- Yamaji, K.: “ Developing E-Infrastructure for Open Science in Japan”, Panel Discussion during Monash Research Data Day, Kuala Lumpur, 2018-10
- Komiyama, Y., Yamaji, K.: “Interdisciplinary research data management service for the whole universities and research institutions in Japan that emphasizes research integrity”,

• 管理基盤

- OSF (OSS) を運用してみたかった潜在機関との連携
 - ドキュメントやインストールコードの整備
- オリジナルに加えて機関向けの機能への評価

• 公開基盤

- 既存のリポジトリシステムの精細な評価資料の評価
- データリポジトリとしてのInvenioの魅力とJAIRO Cloud (機関リポジトリ) を成功例にもつNIIとのコラボレーションに対する評価

今後の開発と運用準備

- 2019年度前半
 - 管理：研究活動サポート機能の強化&運用体制の構築
 - 公開：ワークフロー機能の強化とシステム移行ツールの開発
 - 検索：KBの精細化&複数アルゴリズムの導入

- 2019年度後半



試行運用
ケーススタディ



JAIRO Cloud
移行実験



分野DBからの
データの取込み

連携機能の強化

スケールアウト&運用システムの強化&効率化

- 2020年度：本格運用

RCOS
yamaji@nii.ac.jp