

研究データ管理基盤（GakuNin RDM）

込山悠介

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

トピック

- イントロダクション

- NII RDC（リサーチデータクラウド）のコンセプト

- **GakuNin RDMの概要**

- 研究データ管理サービスGakuNin RDM
- 本サービスの機能の紹介
- 開発ロードマップ

- 実証実験のご案内

- 実験のスケジュール
- 実験への参加方法

イントロダクション

大学で研究データ基盤が必要な理由

研究推進の観点

オープンサイエンス促進

- 研究データが管理されていれば、研究データ共有ができ再現実験や共同研究の促進
- 研究データが利活用しやすい状態で管理されていれば、公開時に他の研究者から研究データを引用
- 研究データの公開により海外研究機関との国際連携などが強化

研究公正の観点

研究不正防止

- 公的研究費で行う研究において、論文の内容を証明するための研究データについては、原則10年、所属機関・部局あるいは研究室で保存する事が推奨
- 資金配分機関が競争的資金のプロジェクトの申請時または採択時に、研究データ計画（DMP: data management plan）の提出を推奨・義務化
- 適切に研究データを集約・保存・整理することは、研究データの紛失や流出を防止

オープンサイエンスに関する国内の政策的経緯

- 2015年3月: **内閣府** 「国際的動向を踏まえたオープンサイエンスに関する検討会」 報告書
→公的研究資金による研究成果の利活用促進の拡大
- 2016年1月 : **政府** 「第5期科学技術基本計画」
→第4章 (2) ③ オープンサイエンスの推進
- 2016年2月: **文部科学省** 科学技術・学術審議会 学術分科会 学術情報委員会
「学術情報のオープン化の推進について」
- 2016年5月: **G7茨城・つくば科学技術大臣会合** つくばコミュニケーション (共同声明)
→オープンサイエンスを効率的に推進し、適切に活用
- 2016年5月: **政府** 「科学技術イノベーション総合戦略2016」
→オープンサイエンスの推進の基本姿勢の下、研究成果・データを共有するプラットフォームの構築
- 2016年7月: **日本学術会議** 「オープンイノベーションに資するオープンサイエンスのあり方に
関する提言」
- 2017年6月 : **政府** 「科学技術イノベーション総合戦略2017」
 - オープンサイエンスの推進: 知識型社会における国際的なルールメイキングへの対応と公的資金によ
る研究成果のオープン化 (データプラットフォームの整備等)
- 2018年6月: **政府** 「統合イノベーション戦略」
 - 知の源泉としてのオープンサイエンス

JSTにおけるオープンサイエンスへの対応 (DMP導入試行をはじめとして), 小賀坂 康志, 第43回SPARC Japanセミナー2016 (2017/2/14) https://www.nii.ac.jp/sparc/event/2016/pdf/20170214_4.pdf (CC BY-ND)

武田英明『研究オープンデータにおける大学と研究者の役割』 (cc BY) より

<https://www.slideshare.net/takeda/ss-106549043>

各分野でのデータ共有の進展

WDC (1950s-)

サイエンス

オープンアクセス

シリアルズ・クライシス

大学図書館/STM出版

インターネット

デジタル化

WWW(Web)

オープンソース

FTP

WWW

1990年代

ICSU WDS (2008)

Budapest Open Access Initiative (2002)

SPARC

arXiv.org
オープンアクセス出版

機関リポジトリ

政府

オープンガバメント

オープンデータ

Web 2.0

Creative Commons

オープンイノベーション

市民科学

オープンサイエンス

科学データ共有

科学データ・アーカイブ

データリポジトリ

研究のオープン化

研究オープンデータ

2000年代

Cloud Computing

2010年代

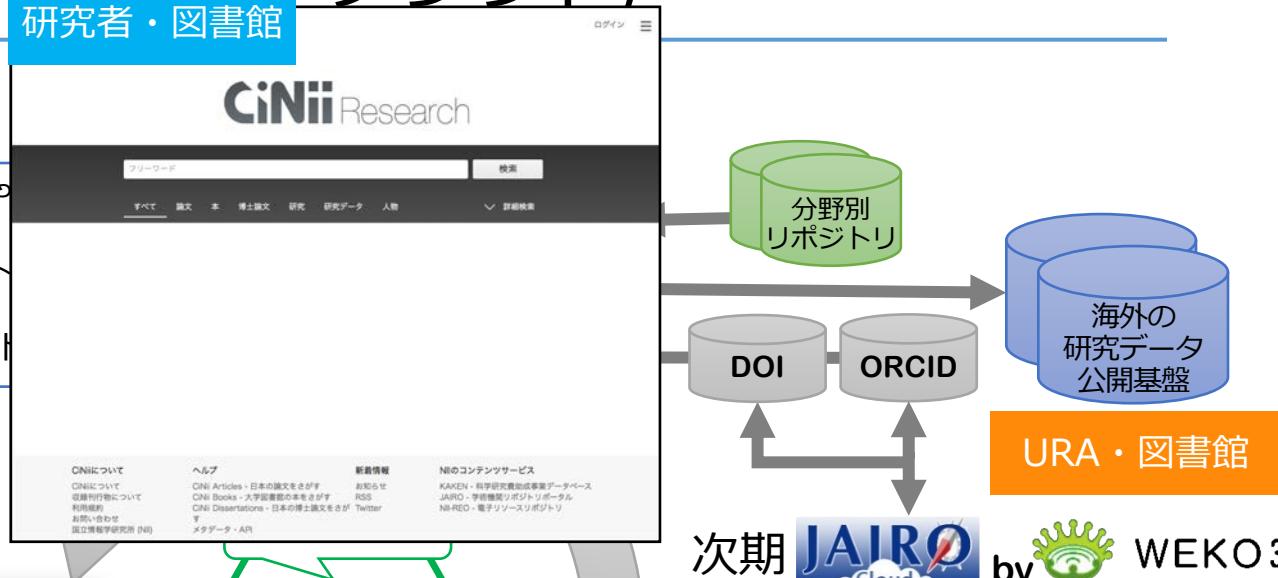
武田英明『研究オープンデータにおける大学と研究者の役割』(cc BY) より

<https://www.slideshare.net/takeda/ss-106549043>

NII RDC (リサーチデータクラウド)

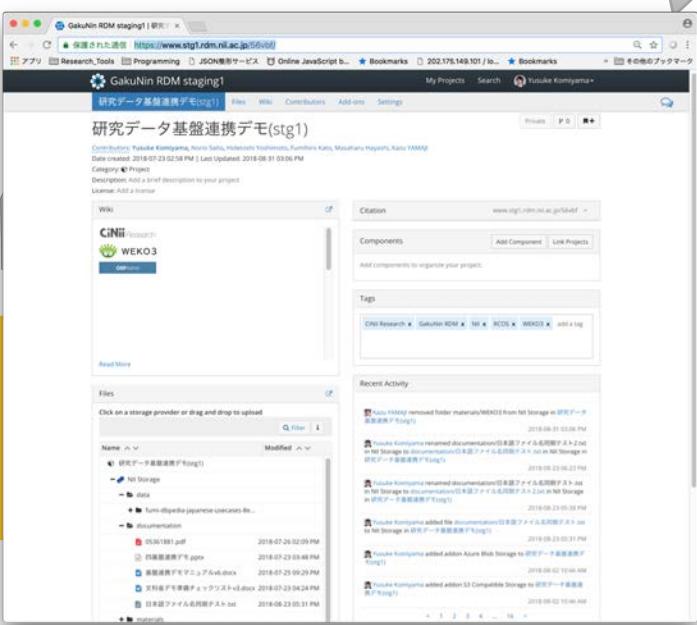
研究者・図書館

- 機関リポジトリ + 分野別リポジトリや
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
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

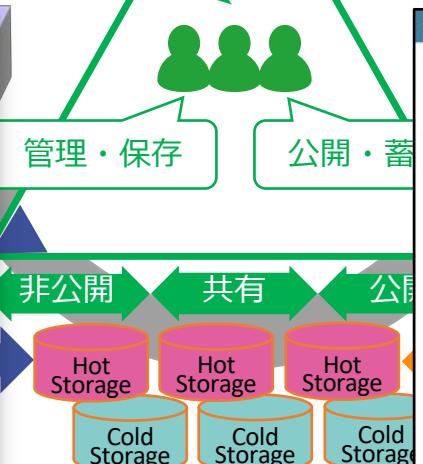


研究者・情報基盤センター

GakuNin RDM



The screenshot shows the GakuNin RDM staging interface. It displays a project titled '研究データ基盤連携デモ(stg1)' with details like creation date (2018-07-23 08:58 PM) and last update (2018-08-31 03:06 PM). The interface includes sections for '研究データ基盤連携デモ(stg1)', '研究データ基盤連携デモ(stg1)', 'Recent Activity', and 'Files'. It shows a file list with items like '研究データ基盤連携デモ(stg1)' and 'Hot Storage'.

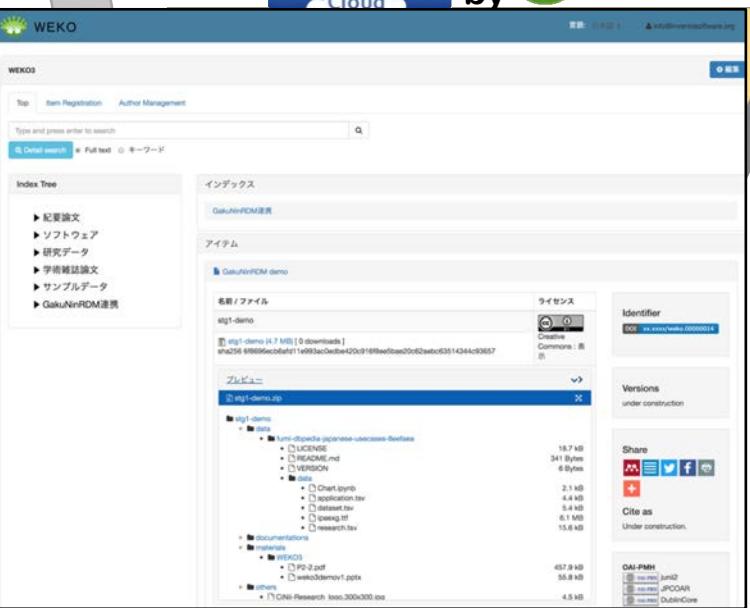


長期保存対応ストレージ領域

 直結
GakuNin Cloud

次期 JAIRo Cloud

by WEKO3



The screenshot shows the WEKO3 interface. It includes sections for 'Top', 'Item Registration', and 'Author Management'. A search bar at the top has 'Full text' selected. The 'Index Tree' section lists categories like '紀要論文', 'ソフトウェア', and '研究データ'. The 'Items' section shows a list of files, including 'alg1-demo' and 'GakuNinRDM demo'. The 'Versions' section is under construction. On the right, there are sections for 'Identifier', 'Share', and 'Cite as'.

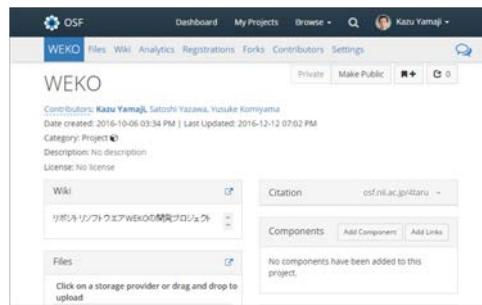
GakuNin RDMの概要

GakuNin RDM とは

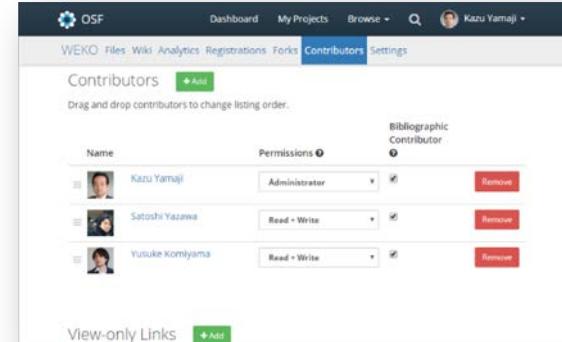
- 学認SPの研究データ管理 (RDM: research data management) サービス
- 主要クラウドプロバイダおよびオンプレミスのオブジェクトストレージの接続が可能
- Webブラウザ上で利用可能な研究データ管理用のストレージサービス
- ストレージ以外にも、研究ツールやソースコードリポジトリ、文献管理ツールの連携が可能

研究データ管理サービス GakuNin RDM

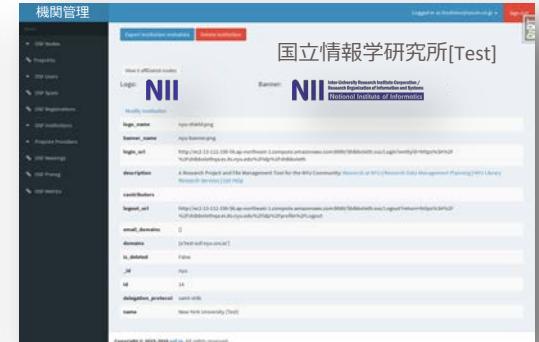
研究プロジェクト単位で
ファイルなどを管理



学認と連携しVO(仮想組織)
メンバーでファイルを共有



機関提供のストレージを利用し
研究証跡を保存・保護



WebアプリケーションはNIIが提供

研究データ管理サービス

機関毎に準備

エクストラストレージ

大学・研究機関毎の既存のクラウドストレージの
事情に合わせてプラグインをカスタマイズ

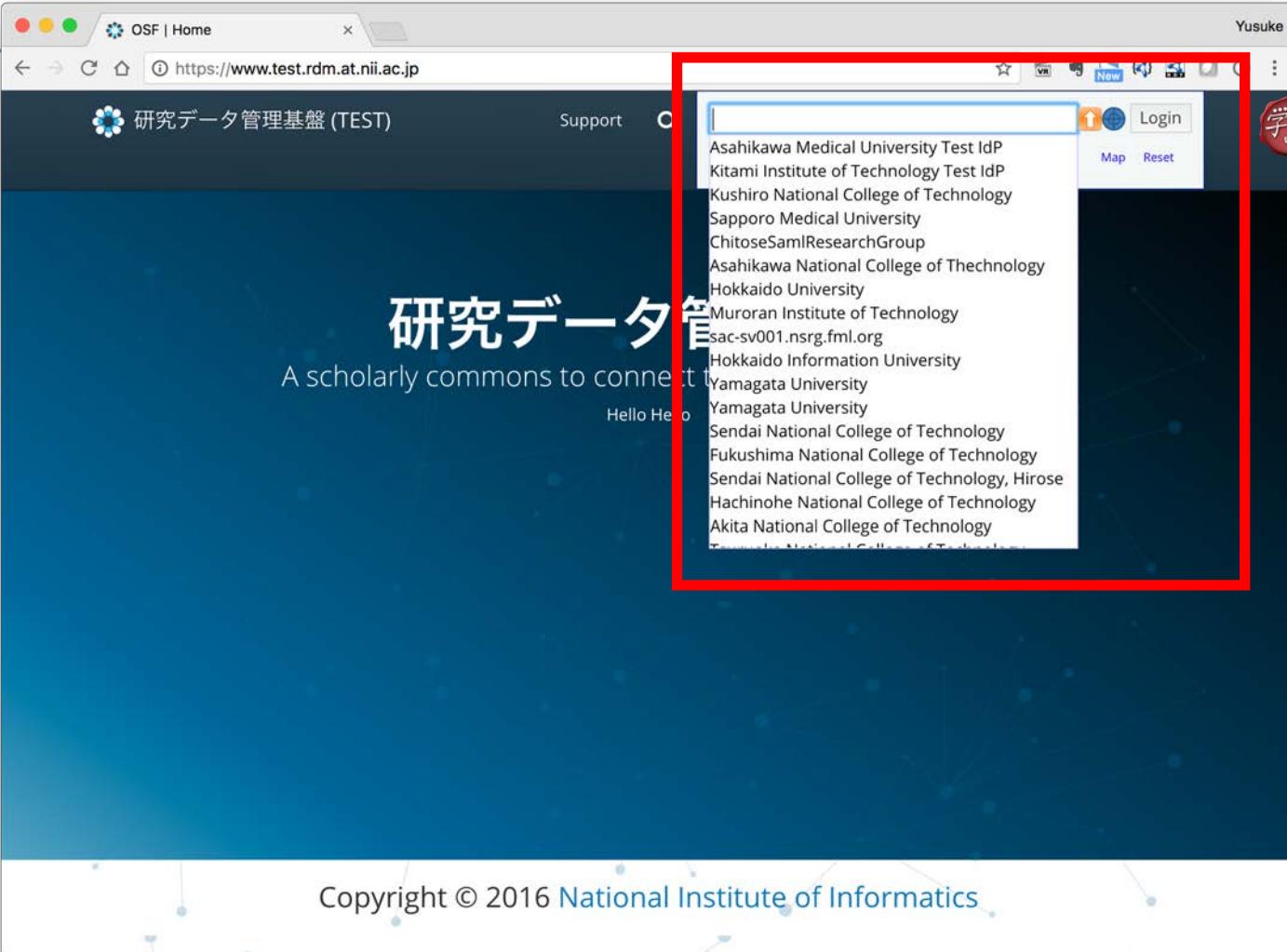
パブリッククラウド
(プロバイダーDC)

プライベートクラウド
(オンプレミス環境)

NIIストレージ
NII提供の最小限の
デフォルト領域
研究者一人当100GB

パブリッククラウド
(プロバイダーDC)

GakuNin RDM (Research Data Management) サービス



The screenshot shows the GakuNin RDM login interface. A red box highlights a dropdown menu listing various academic institutions and research groups that support federated login:

- Asahikawa Medical University Test IdP
- Kitami Institute of Technology Test IdP
- Kushiro National College of Technology
- Sapporo Medical University
- ChitoseSamlResearchGroup
- Asahikawa National College of Thechnology
- Hokkaido University
- Muroran Institute of Technology
- sac-sv001.nsrg.fml.org
- Hokkaido Information University
- Yamagata University
- Yamagata University
- Sendai National College of Technology
- Fukushima National College of Technology
- Sendai National College of Technology, Hirose
- Hachinohe National College of Technology
- Akita National College of Technology

The interface includes a "Login" button, a "Map" link, and a "Reset" link. A red circular badge with the text "学認" (Academic Authentication) is visible in the top right corner.

学認 (学術認証フェデレーション) のDS (discovery service)でログイン。

TiQR を使ったGakuNin RDMの認証



GakuNin RDM
A scholarly commons to connect the entire research cycle

Copyright © 2017 National Institute of Informatics

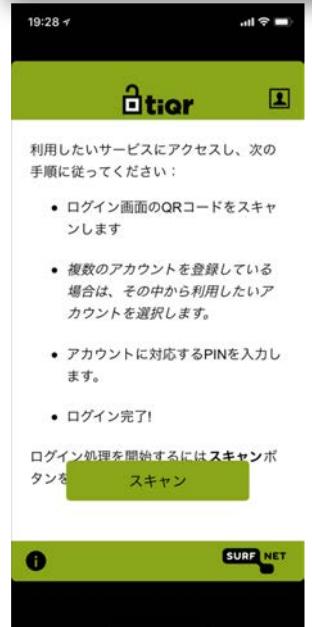


Easy and Safety Login by your Smartphone
Shibboleth Mobile Login by 

Please scan the QR code below by Tiqr application in your smartphone.

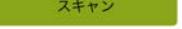


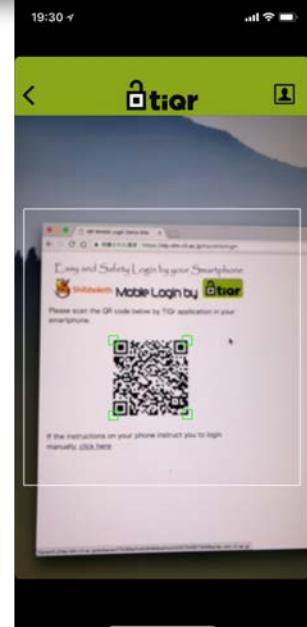
If the instructions on your phone instruct you to login manually, [click here](#)



利用したいサービスにアクセスし、次の手順に従ってください：

- ログイン画面のQRコードをスキャンします
- 複数のアカウントを登録している場合は、その中から利用したいアカウントを選択します。
- アカウントに対応するPINを入力します。
- ログイン完了！

ログイン処理を開始するにはスキャンボタンを 





PINコードを入力してください

ログインするには4桁のPINコードが必要です

●	●	●	●	OK
---	---	---	---	----

1	2 ABC	3 DEF
4 GHI	5 JKL	6 MNO
7 PQRS	8 TUV	9 WXYZ
0		✖



本当にログインしますか？

次のユーザとしてログインします：

フルネーム	test03
ID	test03

ログイン先：

idp.rdm.nii.ac.jp



本サービスの基本機能の紹介

タイトル

共著者

日付

ライセンス

要旨 (Wiki)

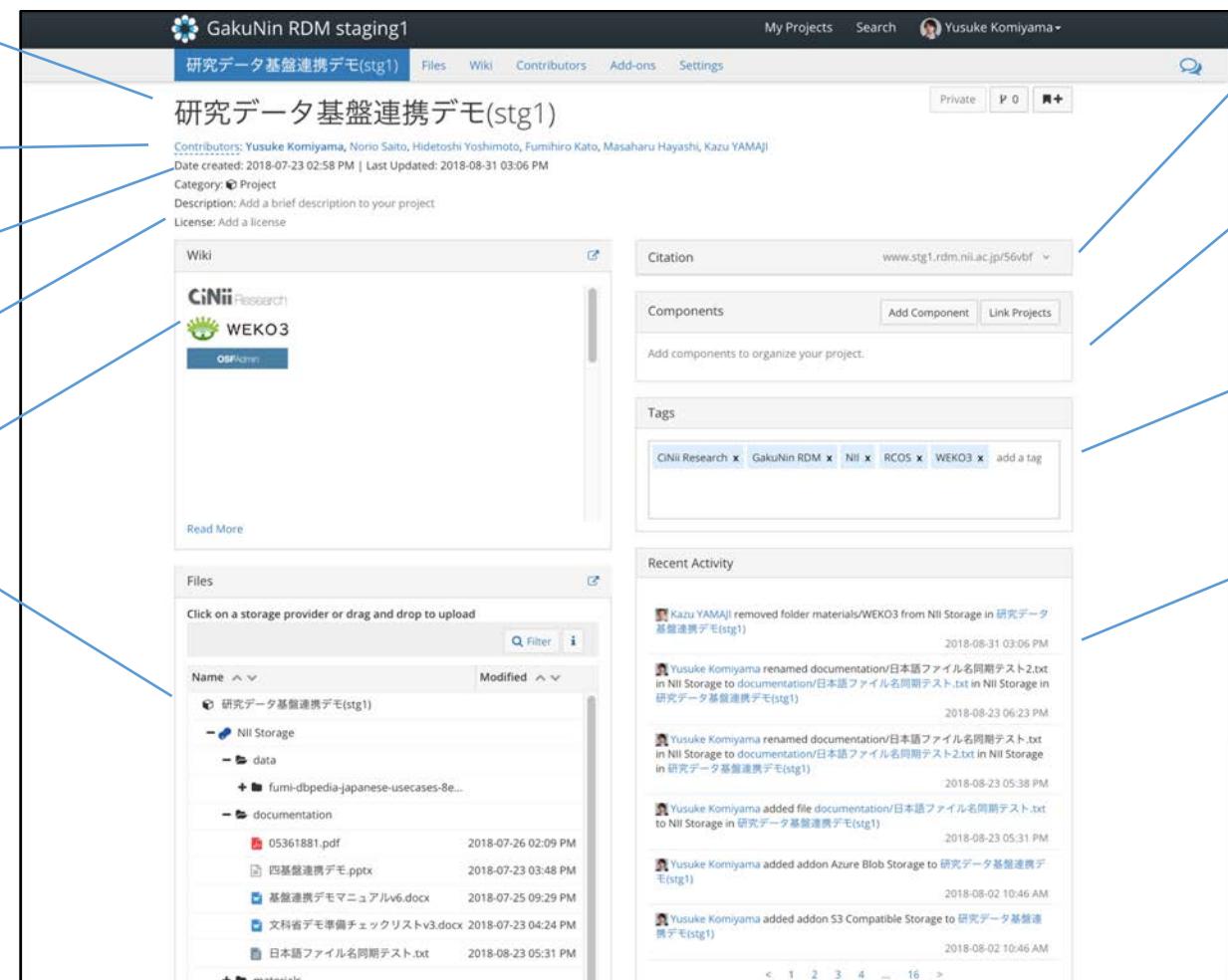
ファイル管理

引用

関連
プロジェクト

検索用タグ

操作ログ



The screenshot shows the GakuNin RDM project homepage for "研究データ基盤連携デモ(stg1)". The page includes sections for Contributors, Description, License, Wiki, Components, Tags, and Recent Activity. Blue boxes and arrows point from the left sidebar to specific elements on the page:

- タイトル (Title) points to the project name "研究データ基盤連携デモ(stg1)"
- 共著者 (Contributors) points to the "Contributors" section
- 日付 (Date) points to the "Date created" and "Last Updated" fields
- ライセンス (License) points to the "License" section
- 要旨 (Wiki) points to the "Wiki" section
- ファイル管理 (File Management) points to the "Files" section
- 引用 (Citation) points to the "Citation" field
- 関連プロジェクト (Related Projects) points to the "Components" and "Link Projects" buttons
- 検索用タグ (Search Tags) points to the "Tags" section
- 操作ログ (Operation Log) points to the "Recent Activity" section

GakuNin RDMで管理した研究プロジェクトのホーム画面

GakuNin RDMと連携可能な外部サービス

・ クラウドストレージ

・ パブリック

-  • Amazon S3
-  • S3互換ストレージ
-  • Azure Blob Storage
-  • Box
-  • Dropbox
-  • Google Drive
-  • One Drive

・ プライベート

-  • ownCloud
-  • NextCloud
-  • OpenStack Swift
 - API ver2, ver3

さくらクラウド
IDCF等

・ データ解析・ワークフロー

-  • JupyterHub
-  • Galaxy
-  • Flowable

・ 図表・スライド共有

-  • Figshare

・ ソースコードレポジトリ

-  • GitHub
-  • BitBucket
-  • GitLab

・ 機関レポジトリ

-  • WEKO2
-  • Dataverse
-  • SWORD互換

・ 文献管理ツール

-  • Mendeley
-  • Zotero

研究者がよく利用するストレージ、データ解析、レポジトリ、文献管理ツールと連携

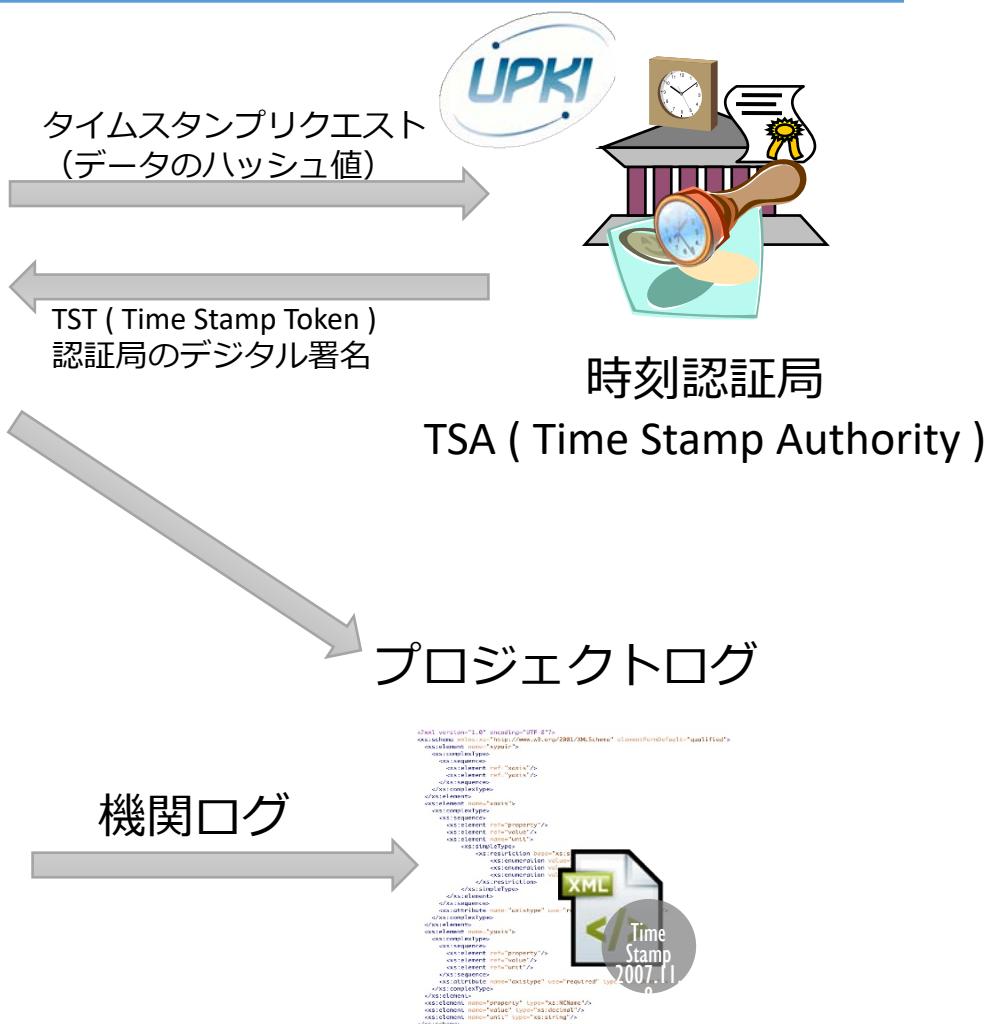


研究証跡保存機能

The screenshot shows the GakuNin RDM interface. At the top, there's a navigation bar with 'Dashboard', 'My Projects', 'Search', and user information. Below it, a file list for 'presentation.pptx (Version: 3)' is shown. The list includes three revisions:

Revision ID	Date	User	Download	MDS	SHA2
3	2017-08-22 09:51 PM	Kazu YAMAJI			
2	2017-08-22 09:50 PM	Kazu YAMAJI			
1	2017-07-06 06:06 PM	Kazu YAMAJI			

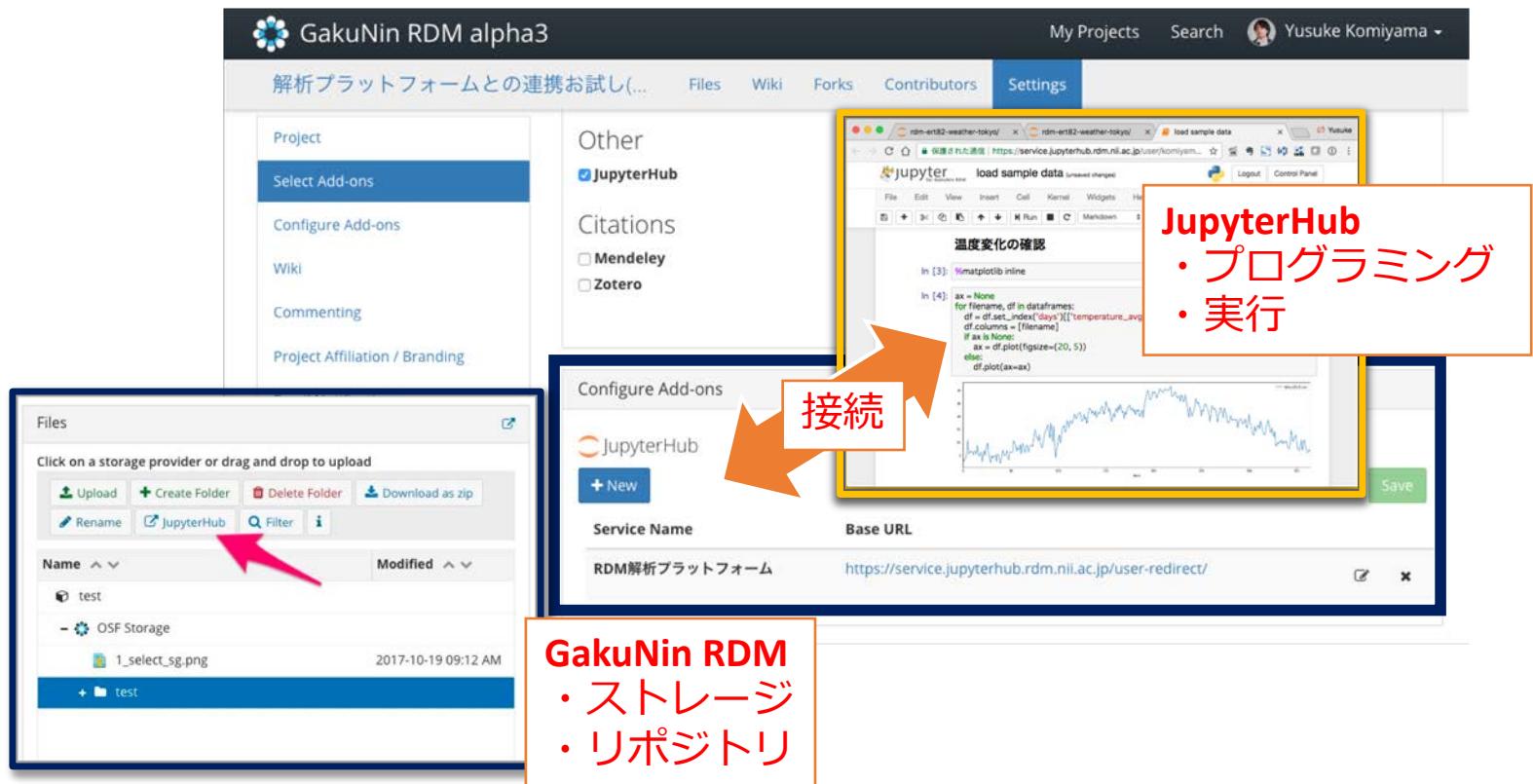
The screenshot shows the 'Institution Management' section of GakuNin RDM. It displays a configuration form for an institution node named 'nyu-shield'. The form includes fields for 'logos_name', 'banner_name', 'login_url', 'description', 'contributors', 'logout_url', 'email_domains', 'domains', and 'is_deleted'. The 'description' field contains a detailed URL for a research project.



データ解析ソフトウェア連携



kubernetes



The screenshot illustrates the integration between GakuNin RDM and a data analysis platform. On the left, the GakuNin RDM interface shows a file storage section with a 'JupyterHub' button highlighted by a pink arrow. In the center, a 'Configure Add-ons' dialog is open, showing the 'JupyterHub' add-on selected. An orange arrow labeled '接続' (Connection) points from this dialog to a JupyterHub session window on the right. The JupyterHub session displays a plot titled '温度変化の確認' (Temperature Change Confirmation) with the following Python code:

```
In [3]: %matplotlib inline
In [4]: ax = None
for filename, df in dataframes:
    df = df.set_index('days')[['temperature_avg']]
    if filename == 'taiwan':
        ax = df.plot(figsize=(20, 5))
    else:
        df.plot(ax=ax)
```

A red box around the JupyterHub session contains the text: **JupyterHub** • プログラミング • 実行.

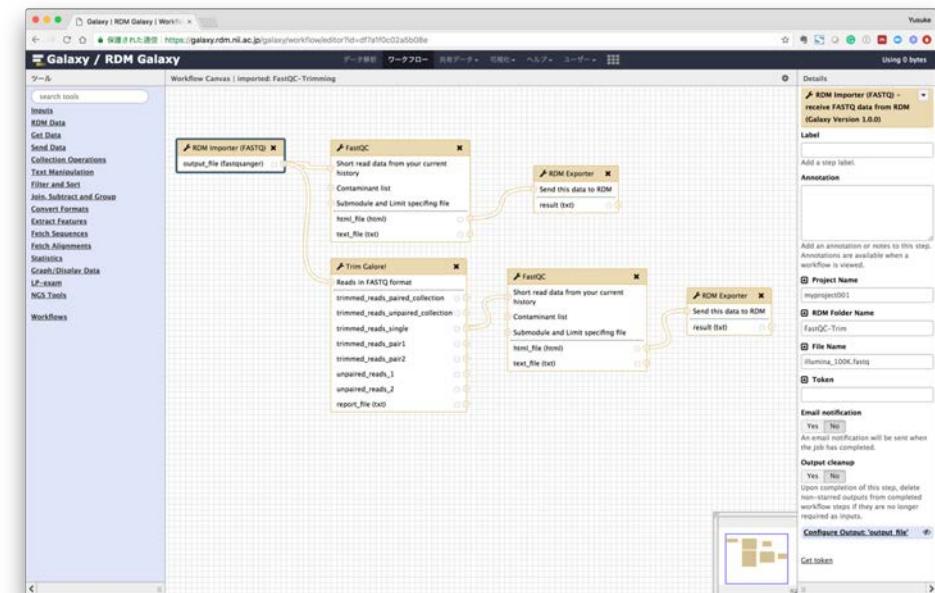
A red box around the GakuNin RDM file storage section contains the text: **GakuNin RDM** • ストレージ • リポジトリ.

- データ解析プラットフォームJupyterHub用のGakuNin RDMアドオンを開発
- 連携によりシステム間でのデータ入出力が容易に
- 双方のアプリケーションでの認証は学認に対応

ワークフローツールとの連携



Galaxyを用いた、科学計算プログラムの自動処理と
GakuNin RDMのストレージの連携

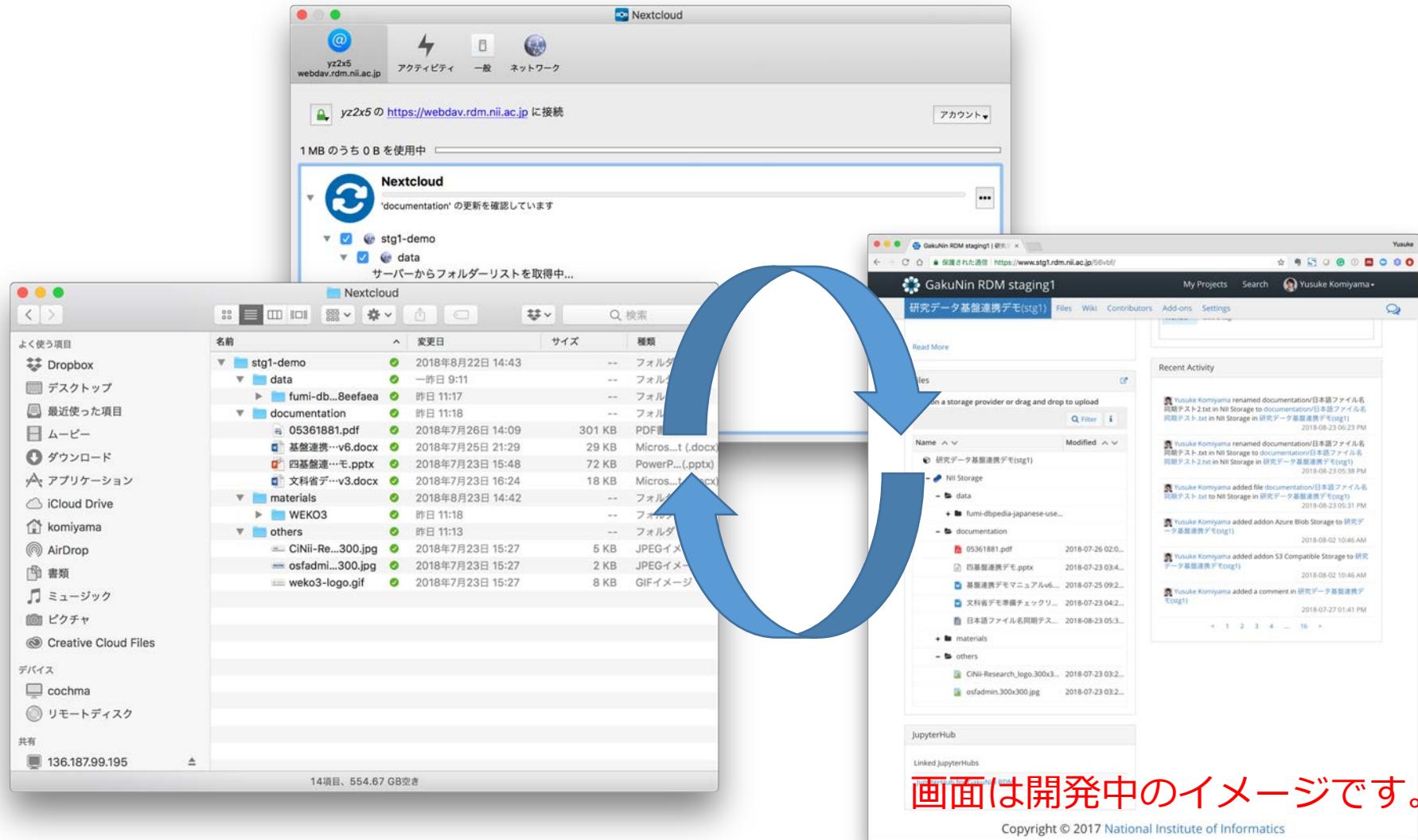


画面は開発中のイメージです。

デスクトップクライアント



Nextcloud ClientのOSSを拡張したGakuNin RDMとデスクトップPCの同期



The diagram illustrates the synchronization process between two interfaces: a desktop client and a web-based application.

Left Side (Desktop Client): A screenshot of the Nextcloud desktop client interface. It shows a sidebar with "よく使う項目" (常用项目) and a main file list. The file list includes a folder named "stg1-demo" containing subfolders "data" and "documentation". Inside "documentation", there are several files like "05361881.pdf", "fumi-db...8eefaea.docx", and "四基盤連携...t.pptx". Other folders include "materials" and "others". The status bar at the bottom shows "14項目、554.67 GB空き".

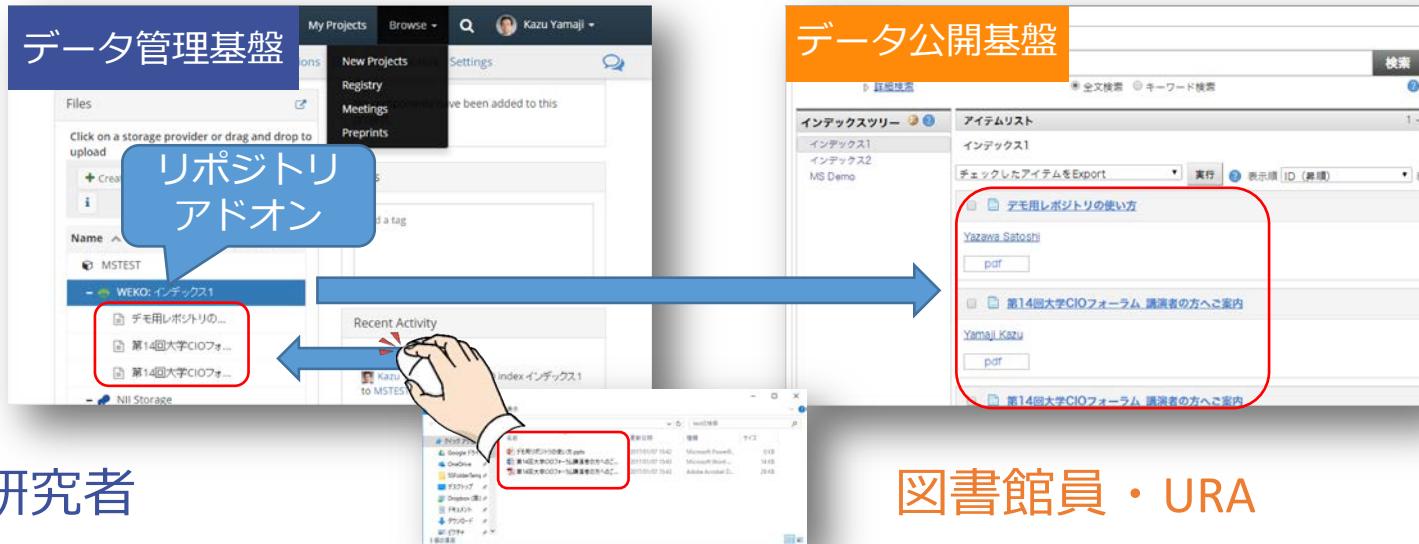
Right Side (Web Application): A screenshot of the GakuNin RDM staging1 web interface. It shows a list of files under "研究データ基盤連携デモ(stg1)". The list includes "研究データ基盤連携デモ(stg1)" (file), "Nil Storage" (directory), "data" (directory), "documentation" (directory), "materials" (directory), and "others" (directory). Inside "documentation", there are files like "05361881.pdf" and "fumi-db...8eefaea.docx". The right side also features a "Recent Activity" sidebar listing various events such as file renames and additions.

A large blue circular arrow icon is positioned between the two interfaces, indicating the bidirectional sync mechanism.

Bottom Right Text: 画面は開発中のイメージです。

Copyright © 2017 National Institute of Informatics

データ管理基盤とデータ公開基盤の役割



研究者

- データ管理機能
- 研究証跡の記録機能
- 公開基盤連携機能
- 非公開データ長期保存機能
- ...

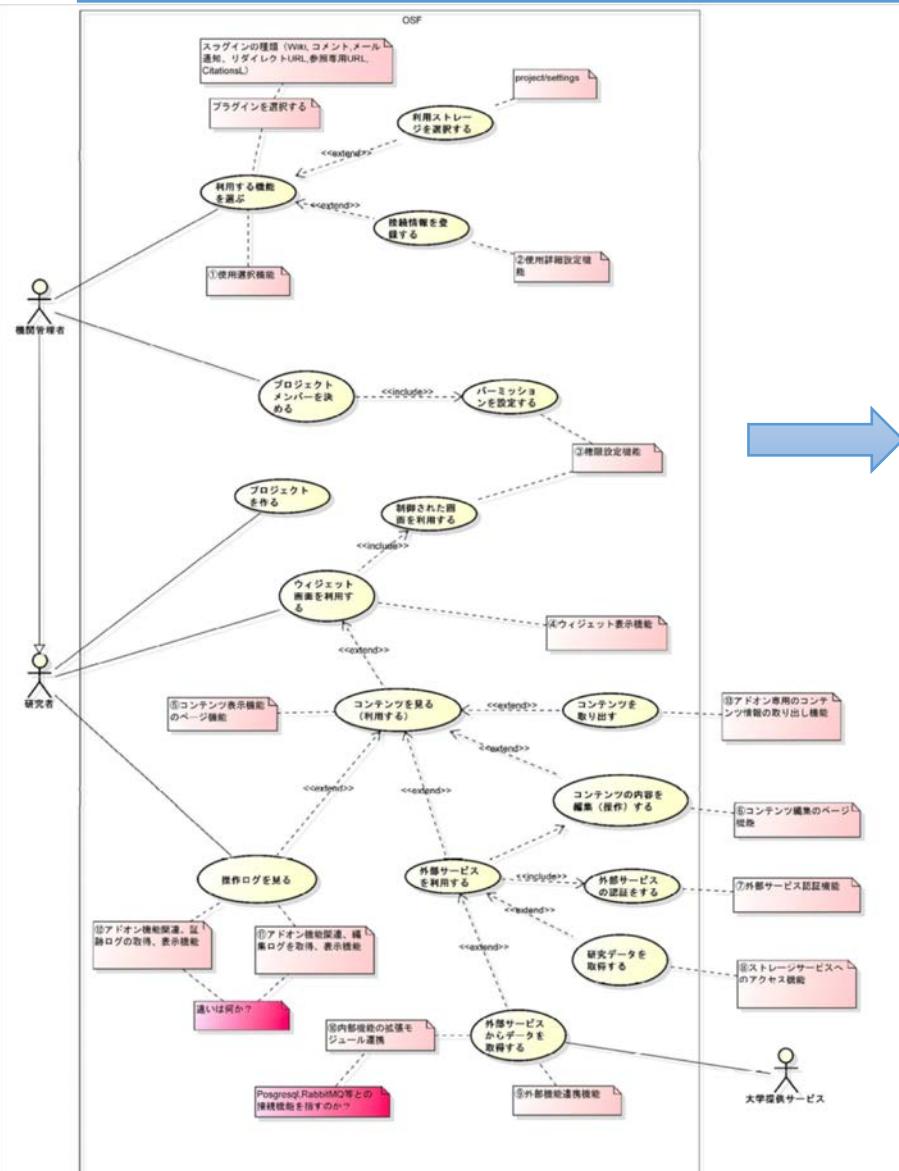
図書館員・URA

- メタデータ管理機能
- データ公開機能
- 管理基盤連携機能
- DOI機能
- ...

大学における業務と整合するように機能を整備



アドオンソフトウェア開発キットの提供



用途

- 開発者向けにユースケースを想定
- GakuNin RDMプラグインSDKを開発



各機関や各研究分野のサービスと接続するために、Pythonのソースコードをテンプレートとして出力

ユースケースの例

- 利用する機能を選択
- 利用ストレージを選択
- 接続情報を登録
- プロジェクトメンバーの決定
- パーミッションの設定
- プロジェクト作成
- ウィジット画面利用
- 制御された画面を利用
- コンテンツ閲覧・利用
- コンテンツ出力
- コンテンツ編集・操作
- 操作ログの閲覧
- 外部サービスの利用
- 外部サービスの認証
- 外部サービスからのデータ取得

開発ロードマップ

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		

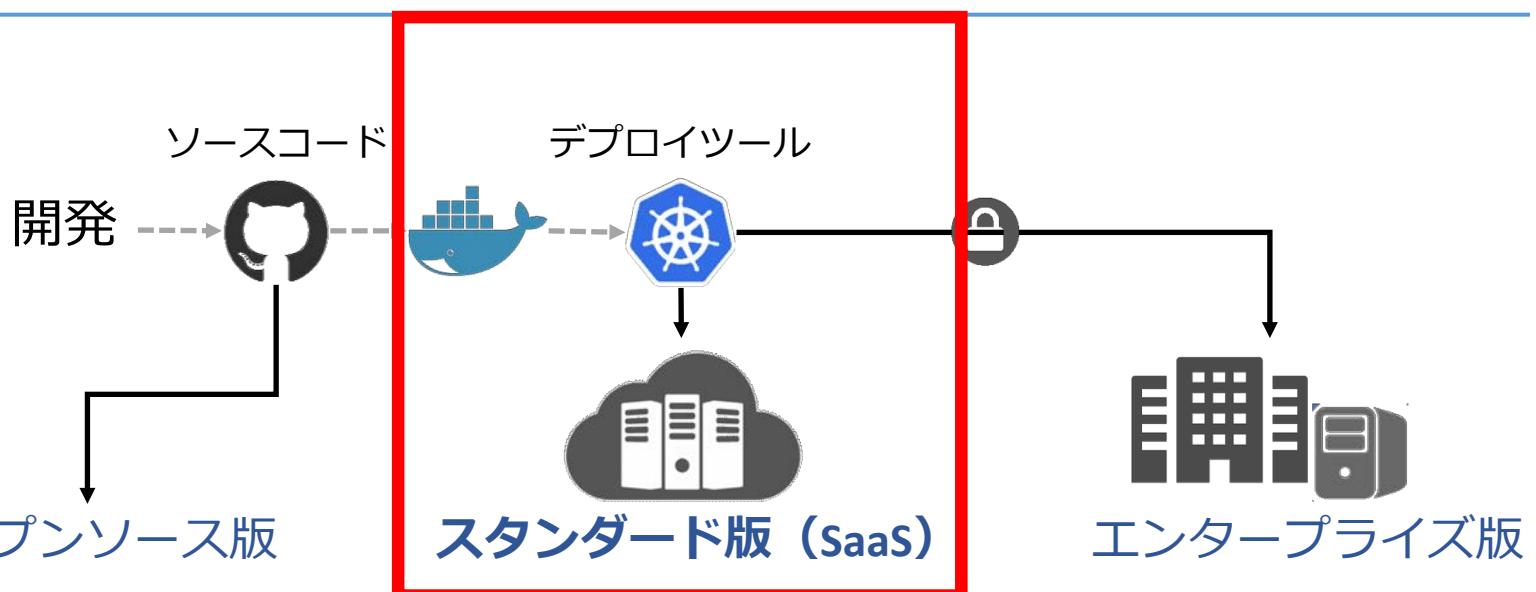
(グレーで「？」の箇所はパートナーが必要なため構想段階)

試作した機能を全国的に提供できるように、各四半期毎にシステムを拡張

実証実験でご意見をいただきながら、
機能追加・性能改善を進めて参ります。



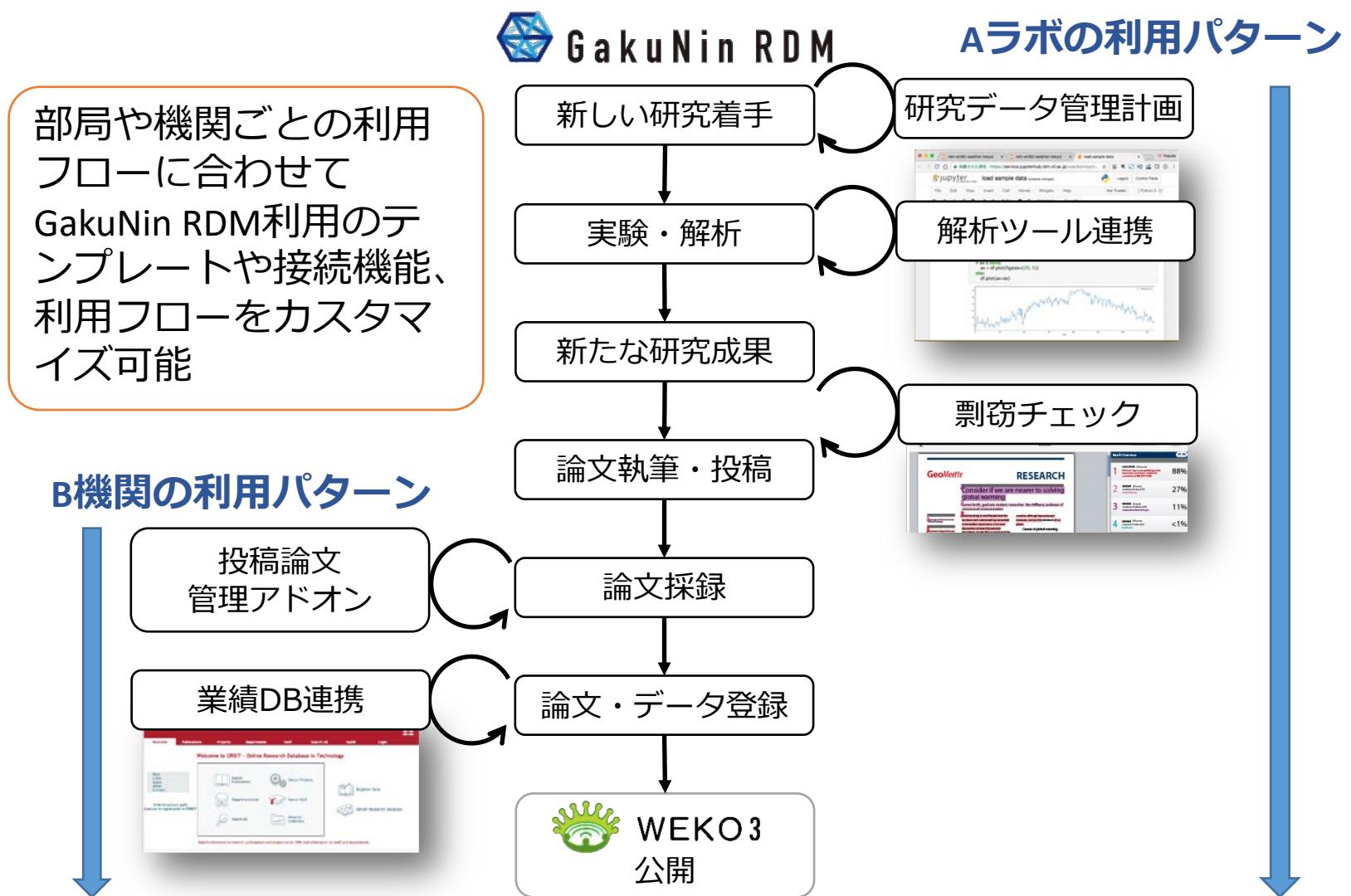
GakuNin RDMの提供方法



- | | | |
|---|--|--|
| <ul style="list-style-type: none">機関が独自にGakuNin RDMを運用したい場合ソースコードとドキュメントを提供運用のノウハウを共有 | <ul style="list-style-type: none">パブリッククラウド上のSaaSとしてNIIが提供機関が用意するストレージを接続して利用可能学認による認証 | <ul style="list-style-type: none">機関のポリシー上スタンダード版ができない場合デプロイコードの提供、あるいはNIIが直接デプロイ機関のIdPによる認証 |
|---|--|--|

機関のポリシー等の制約などにも対応できるように
いくつかの運用オプションを検討中

機関の利用モデルに沿った展開



まとめ

・ 研究者にとってのメリット

- ・ 研究室、共同研究者間で研究データをクラウドストレージで共有
- ・ 研究データ管理だけでなく、研究討議のスペースとしても利用
- ・ 分散している研究データが集約され、データ解析の効率が向上
- ・ 研究費申請書でDMPの項目を書く際の負担が軽減

・ 機関（大学図書館、研究支援部門）にとってのメリット

- ・ 図書館の機関リポジトリと連携しており、研究データ公開が簡単
- ・ 所属研究者のDMPの設計を簡略化し、研究推進を支援
- ・ 研究公正機能（研究証跡の保全、データの不正検出）で所属研究者の研究の透明性・再現性を向上
- ・ 機関の調達事情や情報セキュリティポリシーに合わせたクラウドストレージを選択

- ・ 研究推進のための研究者によるデータ管理・公開を容易にします
- ・ GakuNin RDMは研究に便利なデジタルインフラを提供し、研究公正機能で研究不正を監視して機関のブランドを守ります

実証実験のご案内

実証実験のスケジュール



- | | |
|----------------|-------------------------------|
| 2018年12月 | : GakuNin RDM 実証実験開始 |
| 2018年1~3月 | : 基礎機能の機関導入・先行試用 【目標1,000ユーザ】 |
| 2019年4~6月 | : 第1回機能アップデート |
| 2019年10~12月 | : 第2回機能アップデート |
| 2019年4~2020年9月 | : ロングラン実証実験 【目標10,000ユーザ】 |
| 2020年4~6月 | : 第3回機能アップデート |
| 2020年7~9月 | : 実証実験終了のアナウンス、本稼働の案内 |
| 2020年10月~12月 | : GakuNin RDM 本稼働 |

実証実験への参加方法

実験参加のお申込み・お問い合わせ先

宛先：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

メールアドレス：rcos-office@nii.ac.jp

送信の際 【GakuNin RDM実証実験参加希望・<ご所属>・<氏名>】
と入れてください

GakuNin RDMの実証実験へのご参加、
どうか、よろしくお願ひいたします。



RCOS

<http://rcos.nii.ac.jp/>