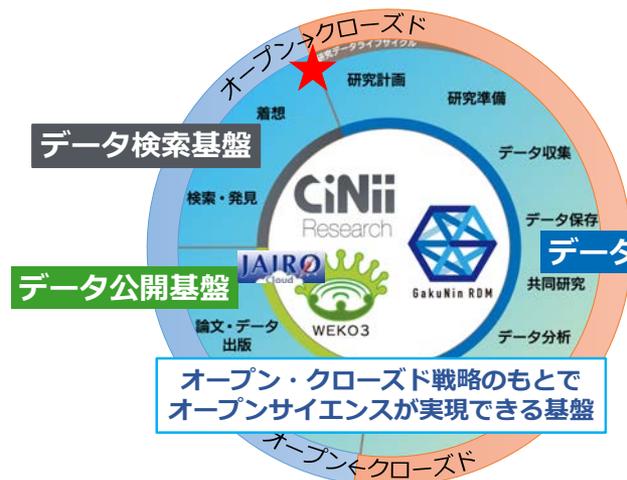


先進的な研究データ管理のための RDMサービスのシステム設計

込山悠介

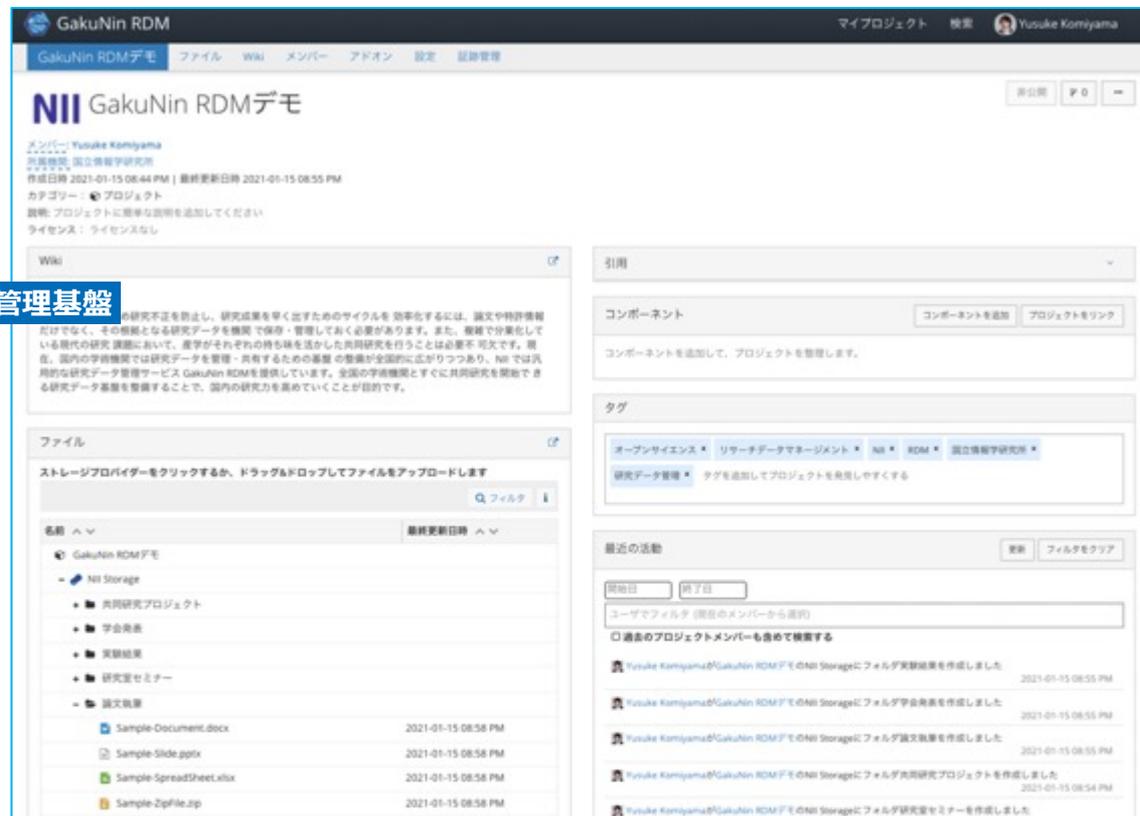
国立情報学研究所

研究データ管理サービスGakuNin RDM



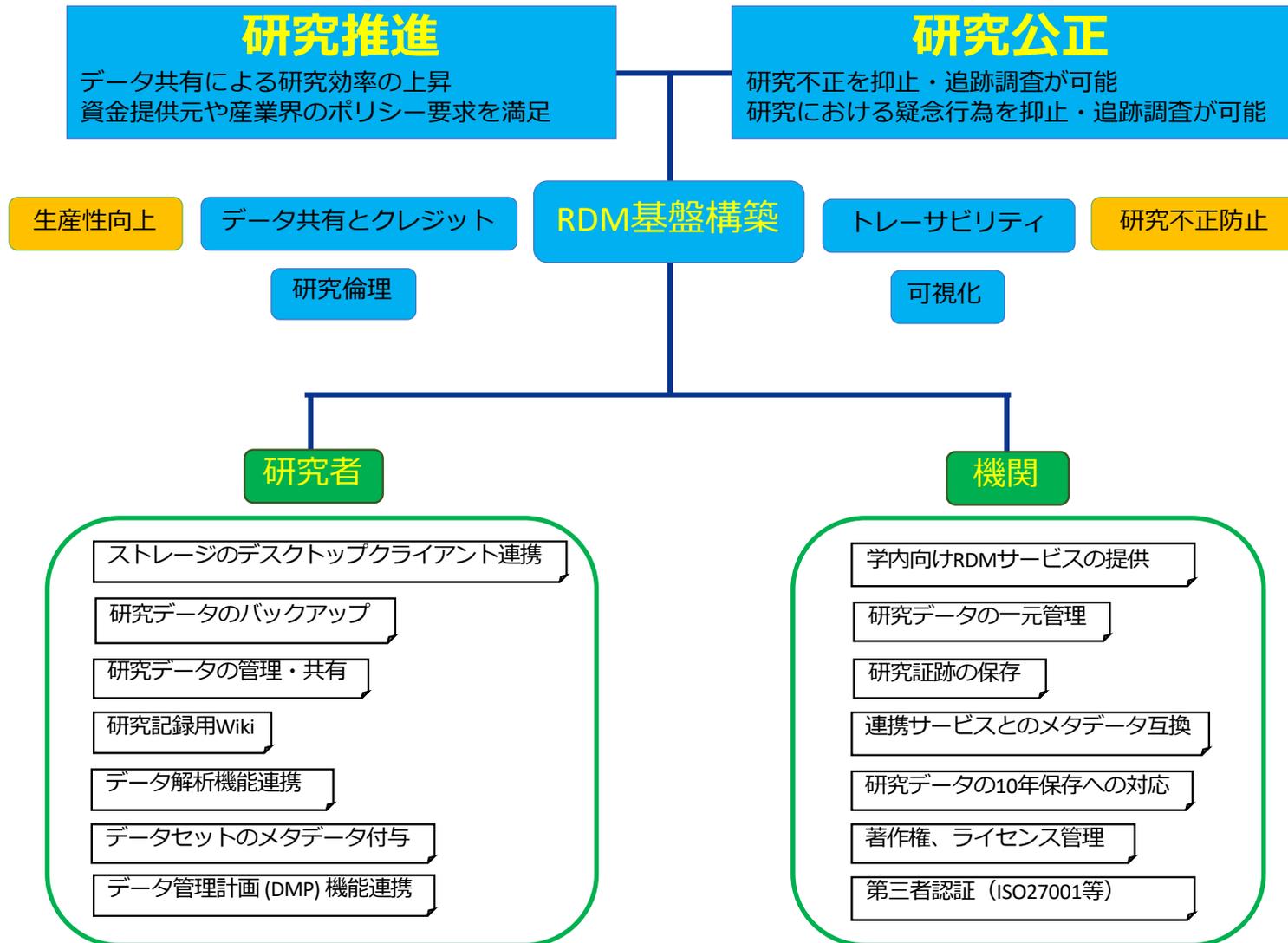
オープン・クローズド戦略のもとで
オープンサイエンスが実現できる基盤

研究データプロセスと
研究データ基盤
NII Research Cloud



GakuNin RDM (GRDM) は、国立情報学研究所 (NII) で運用している研究データ基盤NII Research Cloudの一部で、研究中のクローズド (非公開、制限共有) なデータを取り扱う事に特化した研究データ管理サービス。2021年4月からは24/365運用体制でRDMサービスを全国の学術機関に提供。NIIの他の研究データ基盤とも連携予定。

GakuNin RDMのサービス・ビジョン



Center for Open Scienceが提供する Open Science Frameworkとの関係

- GakuNin RDM (GRDM)は、米国Center for Open Scienceが開発したOpen Science Framework (OSF)のオープンソースソフトウェア版をフォークして開発
- 技術開発に関する情報交換、ソースコードレベルでのバグレポートなど機能改善でCOSと協力
- NIIの研究データ基盤 NII Research Data Cloudでは、公開基盤や検索基盤があるため、サービスのアーキテクチャ上不要なOSFの公開機能については未実装 (リポジトリ機能、公開データの検索機能、カンファレンス支援機能、プレプリント機能等)



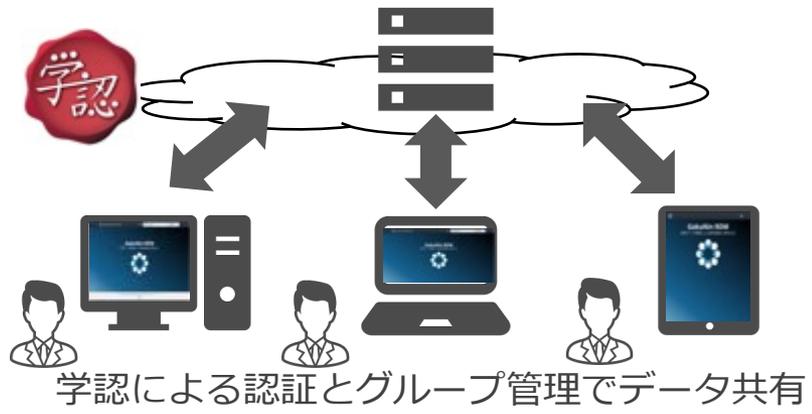
フォーク



GakuNin RDM

GakuNin RDMサービス機能概要

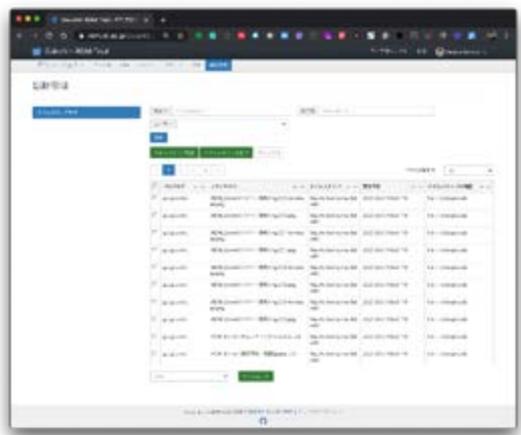
研究データ共有機能（基本機能）



NII研究データ基盤と外部ツールとの連携



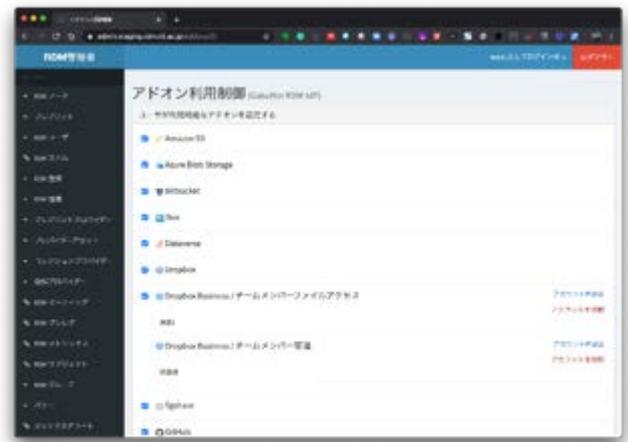
研究証跡の保存機能



時刻認証事業者のタイムスタンプでファイルの存在を証明



機関利用のための管理機能



機関の管理者による組織内のRDMサービスのカスタマイズ

システム設計

GakuNin RDMのストレージ方式

WebアプリはNIIが提供

研究データ管理サービス



機関毎に準備

機関ストレージ

標準ストレージ
(NIIストレージ)

利用開始直後から利用可

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)

S3互換

- ・機関のシステム管理者がいずれかのオブジェクトストレージを学内の標準ストレージとして1種類指定
- ・プロジェクト開始時に自動的にストレージをマウント
- ・NIIストレージは併用不可

機関のシステム管理者が設定すれば全学で利用可

拡張ストレージ

機関毎に準備

拡張ストレージは併用可能



パブリッククラウド
(プロバイダーDC)

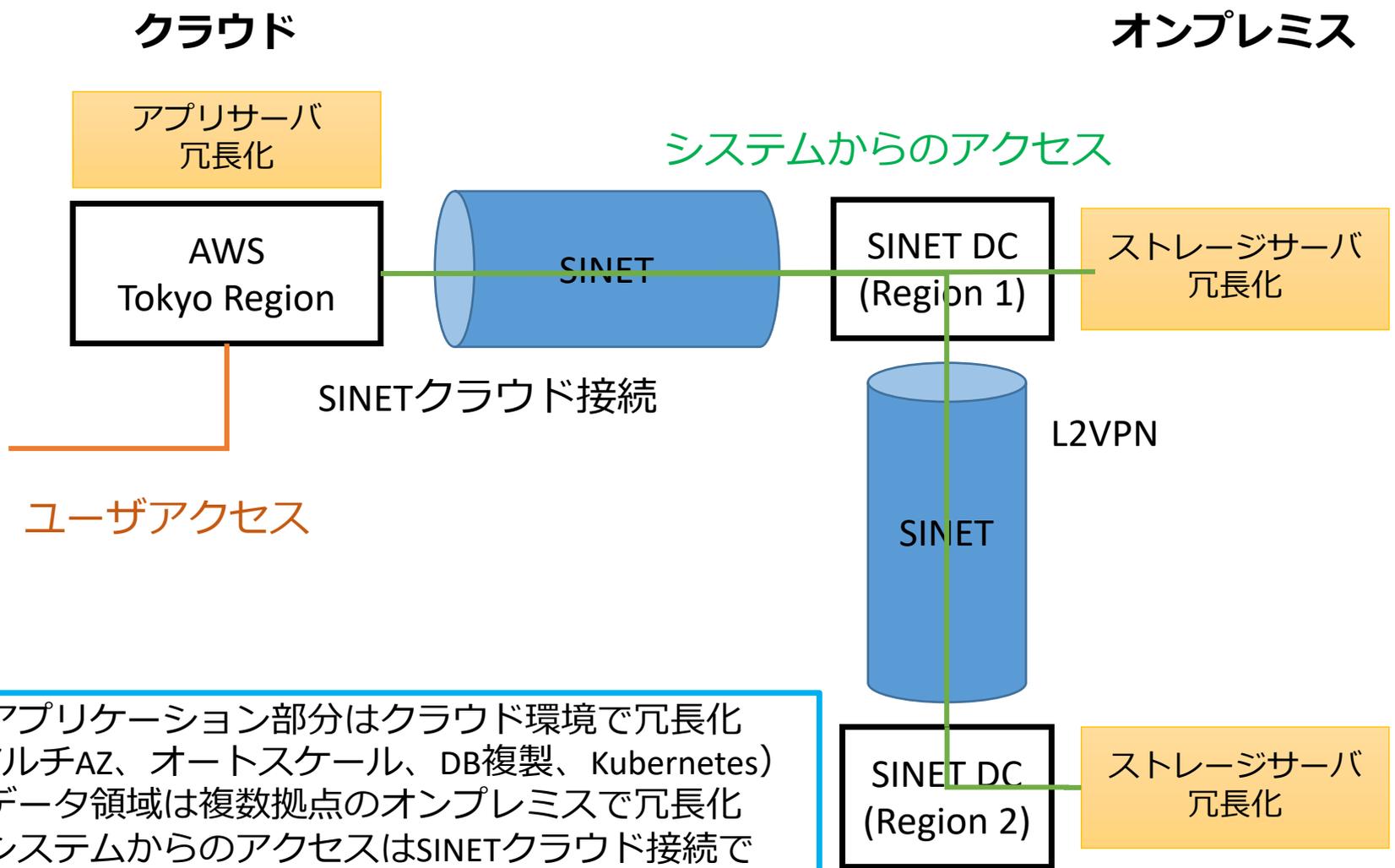


プライベートクラウド
(オンプレミス環境)

S3互換

ユーザ自身で設定すれば利用可能

簡略化したネットワーク図 (標準ストレージ)



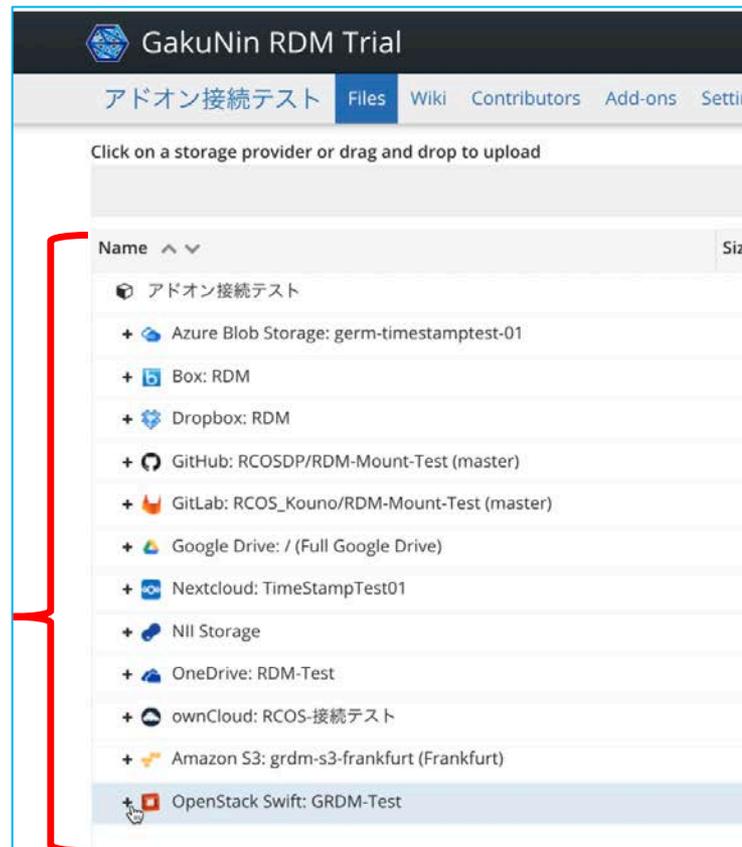
- アプリケーション部分はクラウド環境で冗長化 (マルチAZ、オートスケール、DB複製、Kubernetes)
- データ領域は複数拠点のオンプレミスで冗長化
- システムからのアクセスはSINETクラウド接続でオンプレミスのデータセンターへ直結

GakuNin RDMとクラウドストレージや外部ツールとの連携で研究を促進

NII研究データ基盤と外部ツールとの連携

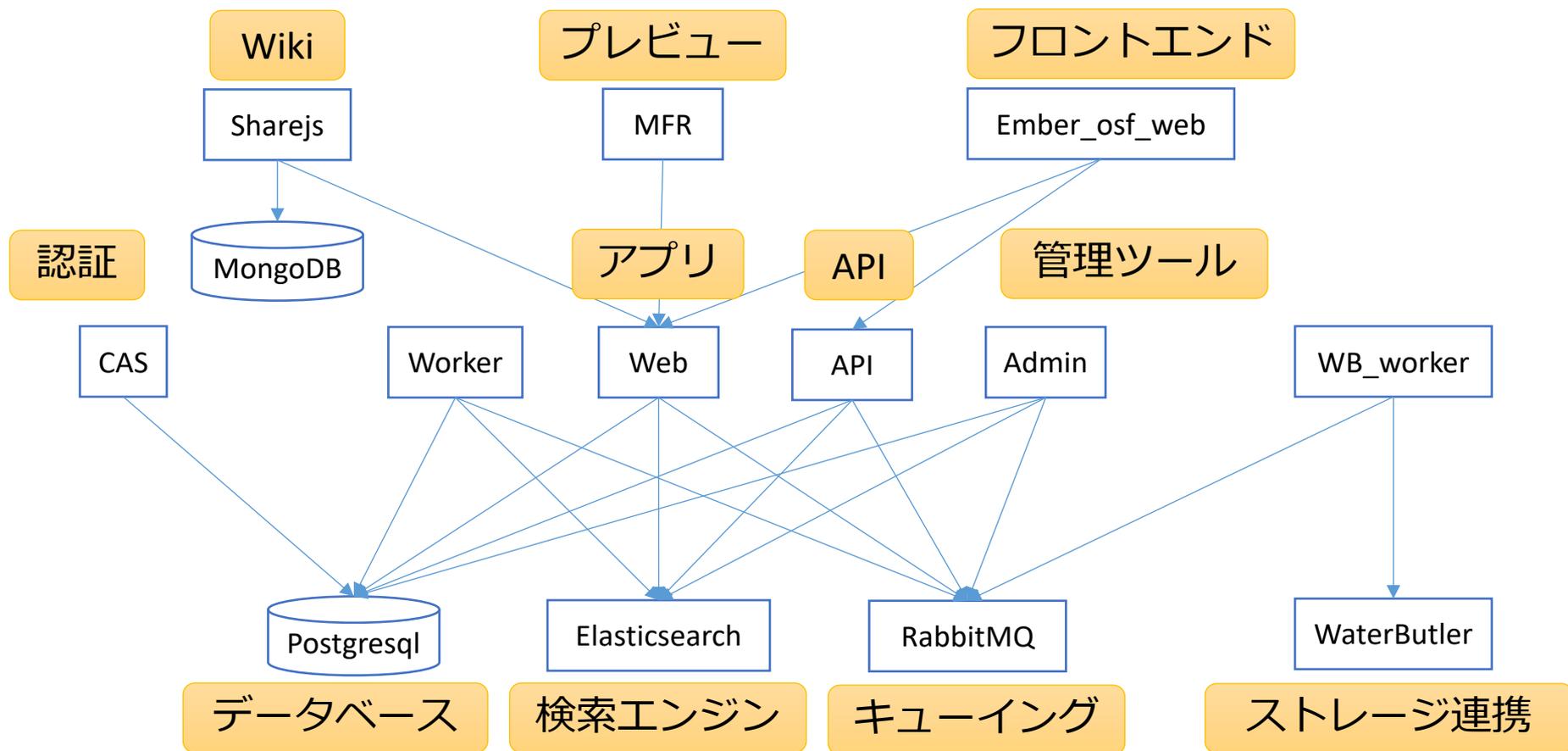


研究プロジェクトに様々な外部クラウドサービスを紐付けて管理



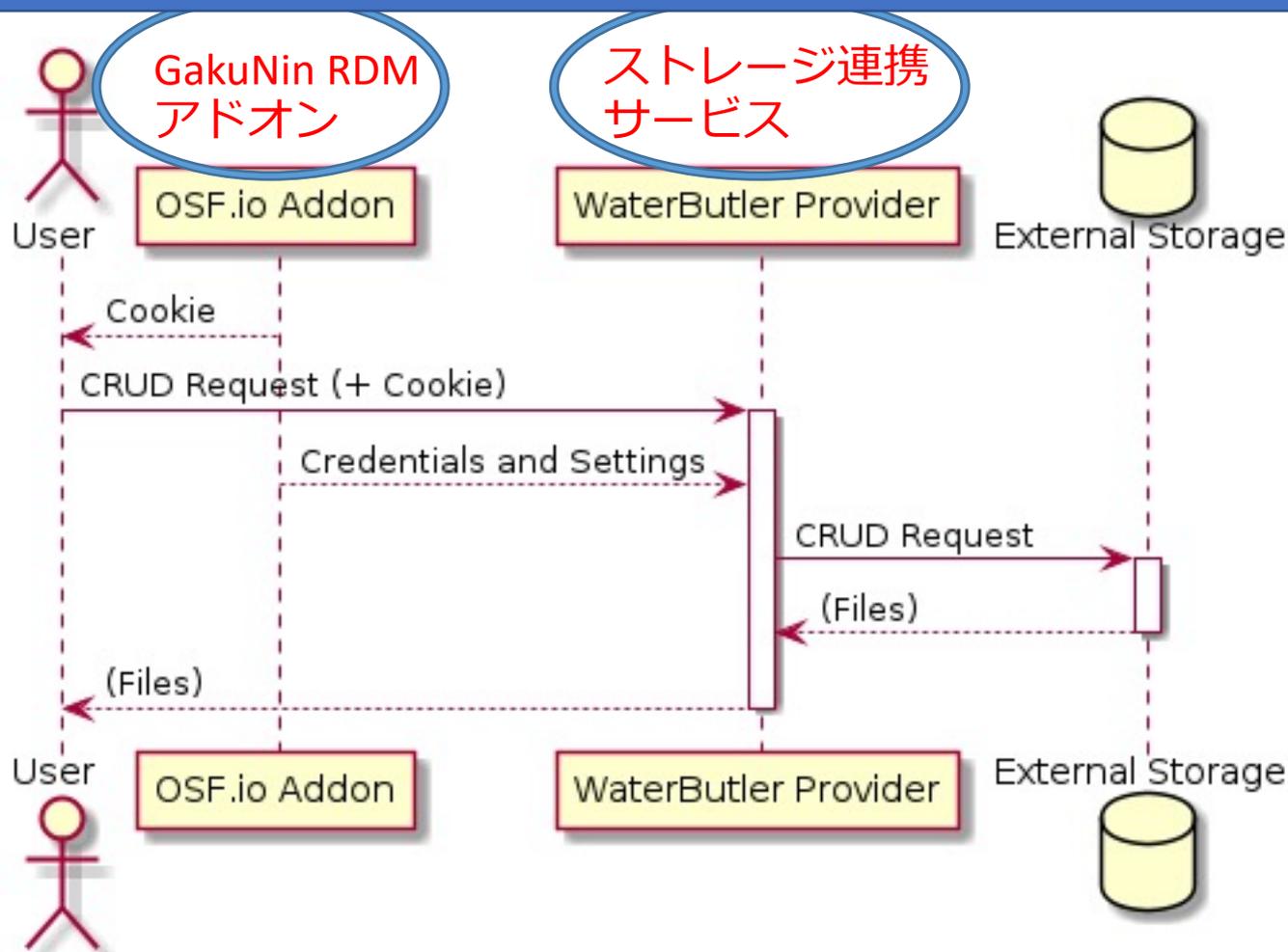
GRDMでは、NII提供分のストレージ以外にも、機関側で契約、所有しているパブリッククラウド、プライベートクラウドのストレージをAPIで接続して利用可能。

GakuNin RDMマイクロサービスアーキテクチャ



GakuNin RDMストレージアドオンにおける外部ストレージとの連携の仕組み

ストレージ用アドオン開発には最低限2つのマイクロサービスへの追加のコードが必要



OSF.io AddonとWaterButler Providerの認証情報の委譲

GakuNin RDMのストレージ連携アドオン

The screenshot displays the 'Add-ons' configuration interface. The top navigation bar includes 'sample', 'Files', 'Wiki', 'Analytics', 'Registrations', 'Contributors', 'Add-ons', 'Settings', and 'Timestamp'. The main content area is titled 'Select Add-ons' and contains the instruction: 'Sync your projects with external services to help stay connected and organized. Select a category and browse the options.'

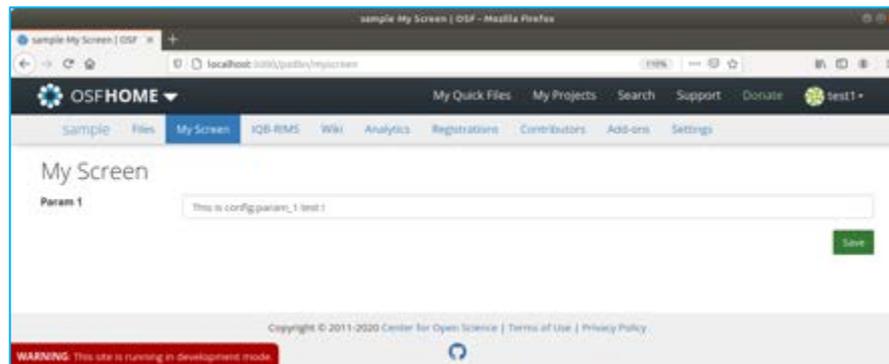
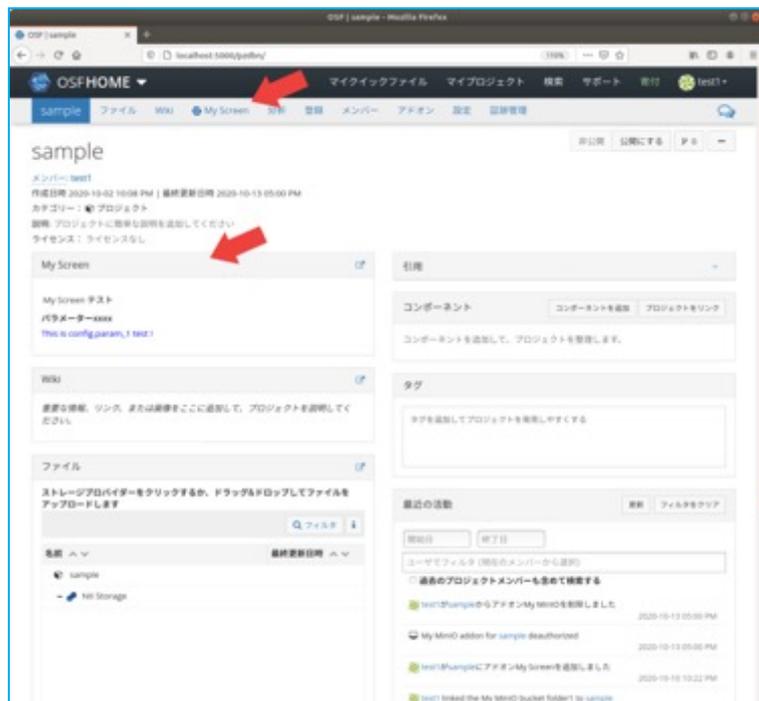
| Categories | Search... |
|------------|----------------------------------|
| All | GitHub Enable |
| Storage | GitLab Enable |
| Other | Google Drive Enable |
| Citations | IQB-RIMS Enable |
| | JupyterHub Enable |
| | Mendeley Enable |
| | My Skelton Disable |
| | Nextcloud Enable |

The 'Configure Add-ons' section for 'My Skelton' is shown below, featuring a text input field containing 'This is test!' and a 'Save' button.

GRDMに外部ストレージ連携することを前提として設計されたアドオン
ストレージ接続やバケット設定に必要な最小限のUIがある



フロントエンドを持つ GakuNin RDMアドオンの設計



フロントエンド (UI) を持つアドオン設計では、バックエンドでのストレージ連携だけでなく、GRDMのプロジェクトのヘッダーのメニューにアドオン名を表示できる。アドオン専用のページを持ち、入力フォームなどを出現させ、プロジェクトのポータルページにモジュールとして追加する。



GakuNin RDMストレージ連携における 認証方式のパターン

- **S3系** (AWS S3, S3互換ストレージ等)
 - 認証方式：hash-based message authentication code (HMAC) 方式。
 - 事前に交換しておいたアクセスキーによって、リクエスト本文のハッシュ値に署名することでリクエストがアクセスキーを持つ者から発行された事を保証
 - AWS S3ではシークレットアクセスキーと年月日、リージョン、AWSサービスからSigning Keyという専用のキーをを作成してリクエストに使用
- **ownCloud系** (ownCloud, Nextcloud等)
 - 認証方式：OAuth認証
 - GRDM側でストレージのアカウント情報を管理
 - 安全のためデバイス単位、アプリ単位のユーザID/パスワードを発行するownCloud (Nextcloud) の拡張プラグインを利用することを推奨している
- **SaaS系** (Dropbox, Google Drive, OneDrive, box等)
 - 認証方式：OAuth認証
 - 各クラウドプロバイダ側でアカウント情報を管理
 - クラウドサービスのアカウント情報で認証を行いGRDMに紐付ける

システム運用

GakuNin RDMの要件やサービスレベル

サービスの性質

- 未発表の研究データを取り扱う専門サービス
- 多くの大学の研究者がデータ共有で利用
- 多様な研究分野のツールと連携が目標
- 実験装置とデータ解析環境とも連携
- 海外出張中や国際共同研究でもストレージ利用
- COVID-19影響による研究者のテレワーク需要拡大
- 高い冗長性、可用性、機密性を要求

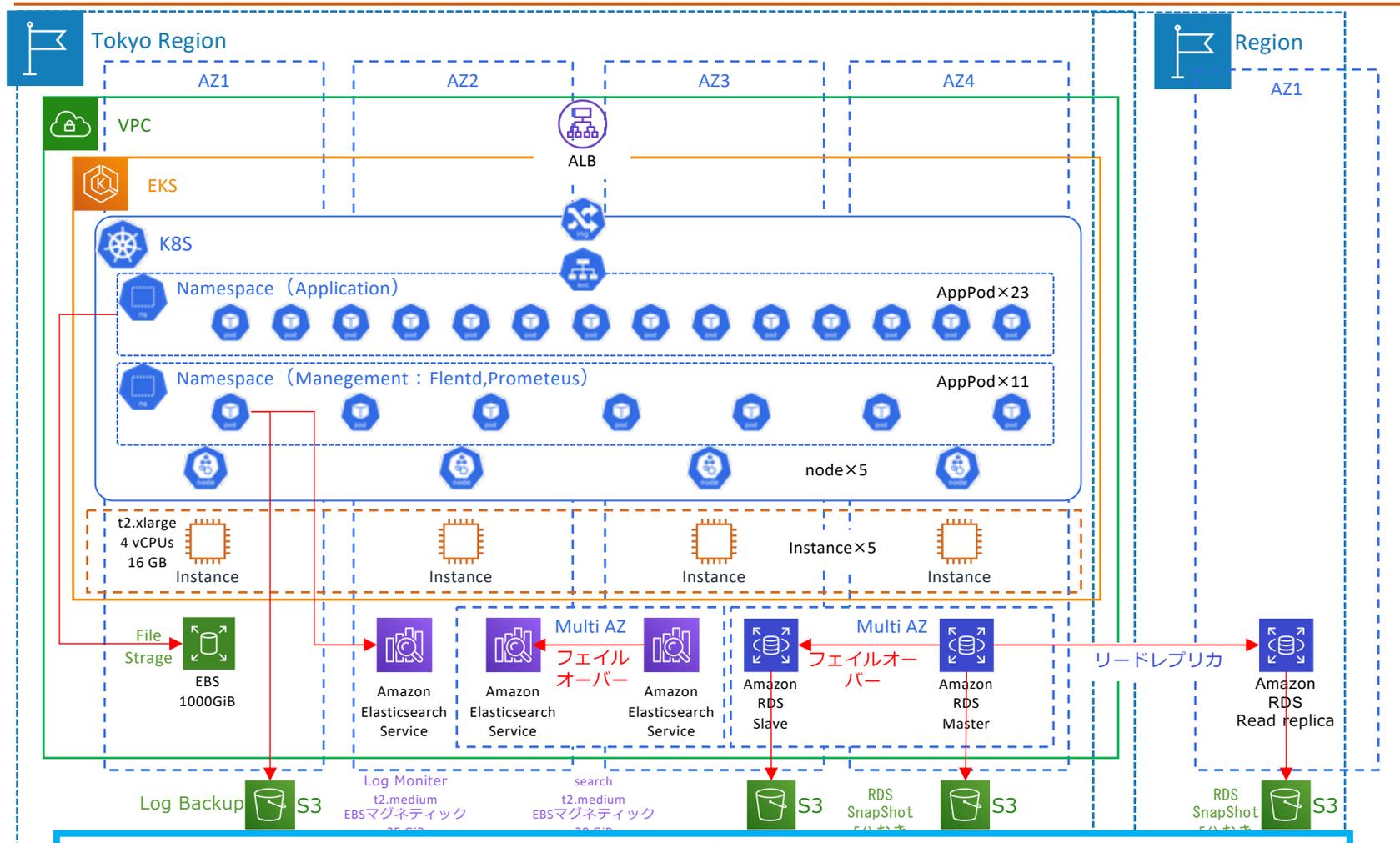
サービスレベル

- 24/365運用を目指す (サービス開始当初は昼間/365)
- 99.5% (サービス障害停止の許容時間3.6H/月)
- 複数データセンター間での複製や高頻度バックアップ
- 無停止でのローリングアップデート等が必要

GakuNin RDM運用で用いているCI/CDツール

- **GitHub:** ソースコードリポジトリ。バージョン管理、コードレビュー
- **TravisCI:** GitHubなどと連携可能な自動単体テスト実行ツール
- **Jenkins:** コンテナのビルドやオーケストレーションの管理ツール、自動デプロイ用ツール
- **Docker:** コンテナ仮想化。アプリケーションの配置・実行用
- **Kubernetes**とはコンテナのオーケストレーションシステム
- **Prometheus:** システムモニタリング、アラート。
Kubernetes上のコンテナの監視
- **Grafana:** 監視プラットフォームのダッシュボード
- **Kibana:** ログ分析プラットフォームのダッシュボード。ログ収集はFluentd, Elasticsearch構成

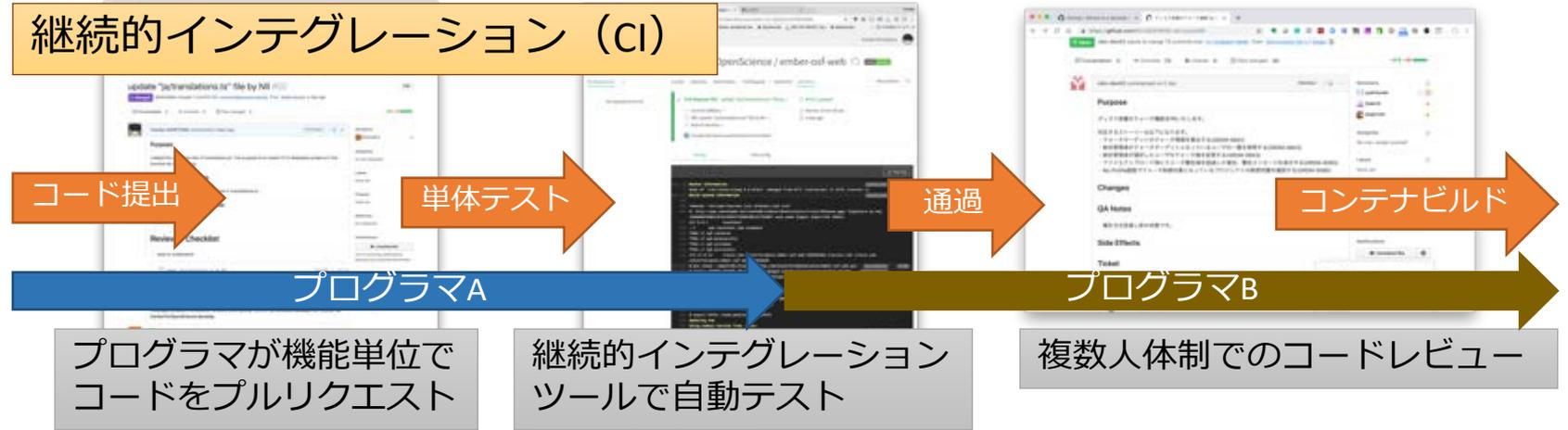
KubernetesによるGakuNin RDMのシステム構成



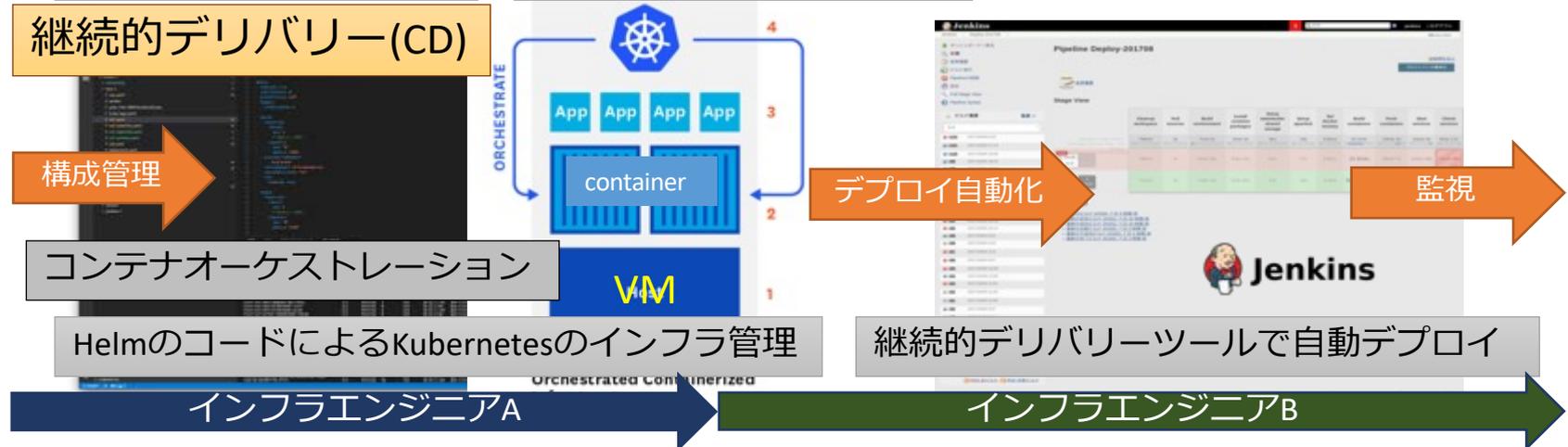
Kubernetesとはコンテナのオーケストレーションシステム。
 アプリ用コンテナのデプロイ、スケーリング、管理のために利用。
 アプリの無停止でのアップデート、サーバ自動復旧などが容易。

サービス運用手法の改善でサービスレベル向上

継続的インテグレーション (CI)



継続的デリバリー (CD)



- ・クラウドネイティブのシステム構成で無停止での機能追加・サーバの冗長化
- ・コードでのインフラ管理による運用・監視自動化、ヒューマンエラー対策
- ・チャットボット導入による窓口業務、サポートの自動化

ユースケース

GakuNin RDM利用機関数



『GakuNin RDM 利用機関数』

<https://meatwiki.nii.ac.jp/confluence/pages/viewpage.action?pageId=41588114>

NII の YouTube チャンネル

GakuNin RDM 事例紹介動画リスト

The screenshot shows the YouTube channel page for GakuNin RDM. The main video player displays a video titled "GakuNin RDM事例紹介動画" (GakuNin RDM Example Introduction Video) with 3 videos, 47 views, and a last update date of 2021/05/23. The video description explains that GakuNin RDM is a research data management base for managing research data and related materials. It highlights features like version management, access control, and research evidence recording. The URL <https://rcos.nii.ac.jp/service/rdm/> is provided.

Below the main video, there is a list of three example introduction videos:

- 1 画像・データ・論文の証跡管理を徹底 - 研究公正 - GakuNin RDM事例紹介動画
国立情報学研究所 - National Institute of Informatics (7:54)
- 2 共同研究におけるデータ共有 - 研究推進 - GakuNin RDM事例紹介動画
国立情報学研究所 - National Institute of Informatics (6:43)
- 3 データ管理からイノベーション創出へ - 大学間連携 - GakuNin RDM事例紹介動画
国立情報学研究所 - National Institute of Informatics (7:35)

A QR code is located in the bottom right corner of the page.

『NIIのYouTube チャンネル GakuNin RDM事例紹介動画リスト』
<https://www.youtube.com/playlist?list=PLKzfdUM-Rtgyj44t821KN1xO5lt78Ea4O>

基盤開発者やサービス提供者の皆様へ

- GakuNin RDMはOSSでGitHub上で開発を行っており Apache License 2.0 で提供。学術機関やクラウドプロバイダー等との共同開発実績あり。
- GakuNin RDMのアドオンソフトウェアを開発・提供される場合、GitHubでプルリクエストいただければコードレビューの上、取り込んで全国へ提供も可能。
- 汎用性の高いOSSのツール、専門分野で広く使われており必要不可欠なツール、国際標準規格に対応したデータ取得API、公開可能なライセンスであることなどの条件をクリアすればGakuNin RDMアドオン化が可能。
- データ基盤間の連携開発、クラウドサービスのAPI連携開発などのご要望がございましたらお問合せ下さい。

GakuNin RDMに関するお問合せ

- GakuNin RDMは利用申請の申込み受付中です。まずは貴機関の情報基盤センターまたは情報システム部門のシステム導入ご担当者様にご相談ください。
- 申込みはGakuNin RDMユーザサポートページのWeb申請フォームにて、貴機関のシステム担当部門のご担当者様から、お申し込みしていただく必要がございます。
 - GakuNin RDMユーザサポートページ：
<https://meatwiki.nii.ac.jp/confluence/display/gakuninrdmusers>
- ご不明点・ご質問は下記のメールアドレスまでお問合せ下さい。
 - GakuNin RDMユーザサポート
rdm_support@nii.ac.jp

