

# 世界的な研究データ管理の方向性と GakuNin RDMの将来構想

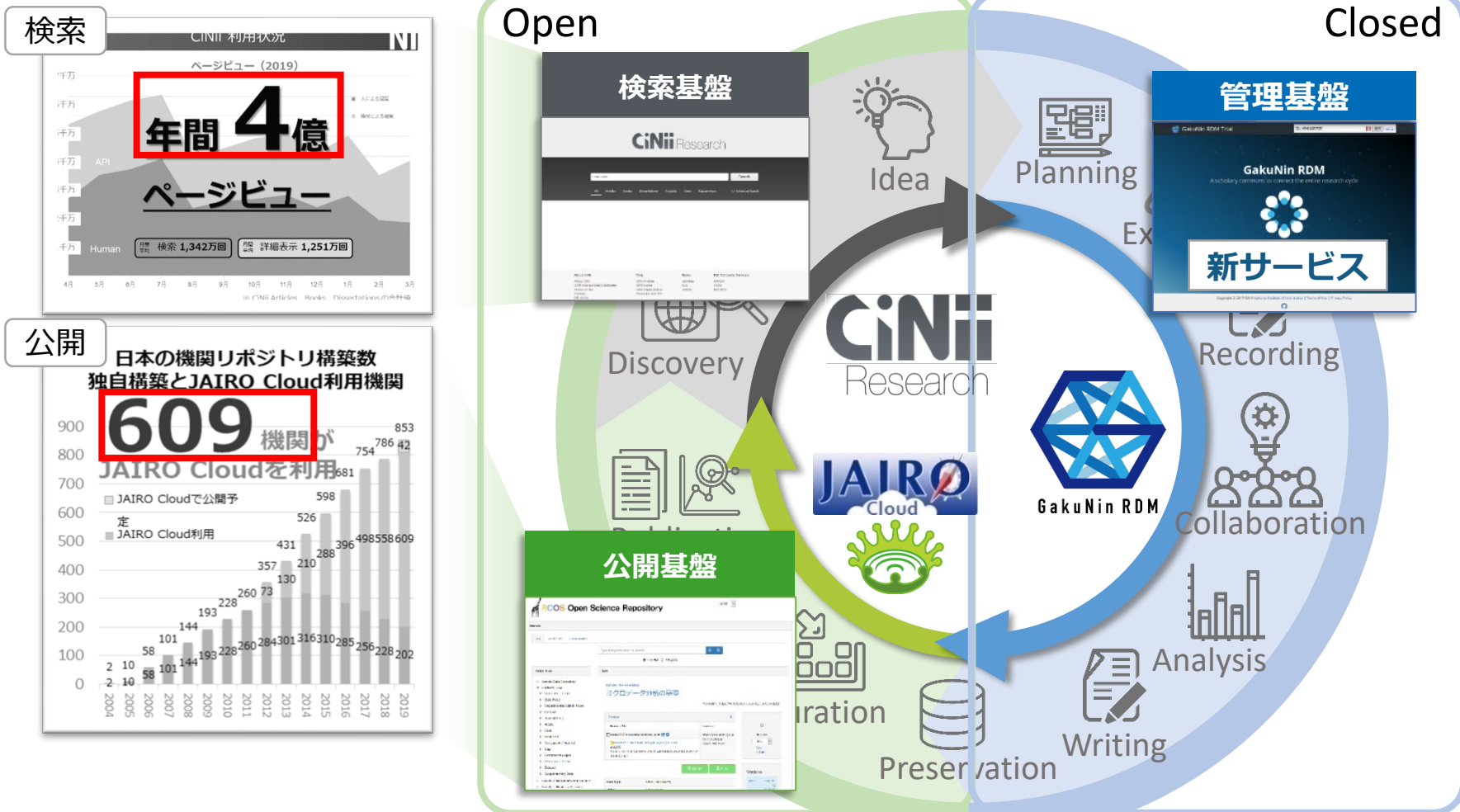
国立情報学研究所 山地一禎

第4回京都大学研究データマネジメントワークショップ

日時：2020年9月19日(土) 11時30分～17時00分

# NII Research Data Cloud

現在のサービス → オープン・クローズ戦略を支援する基盤

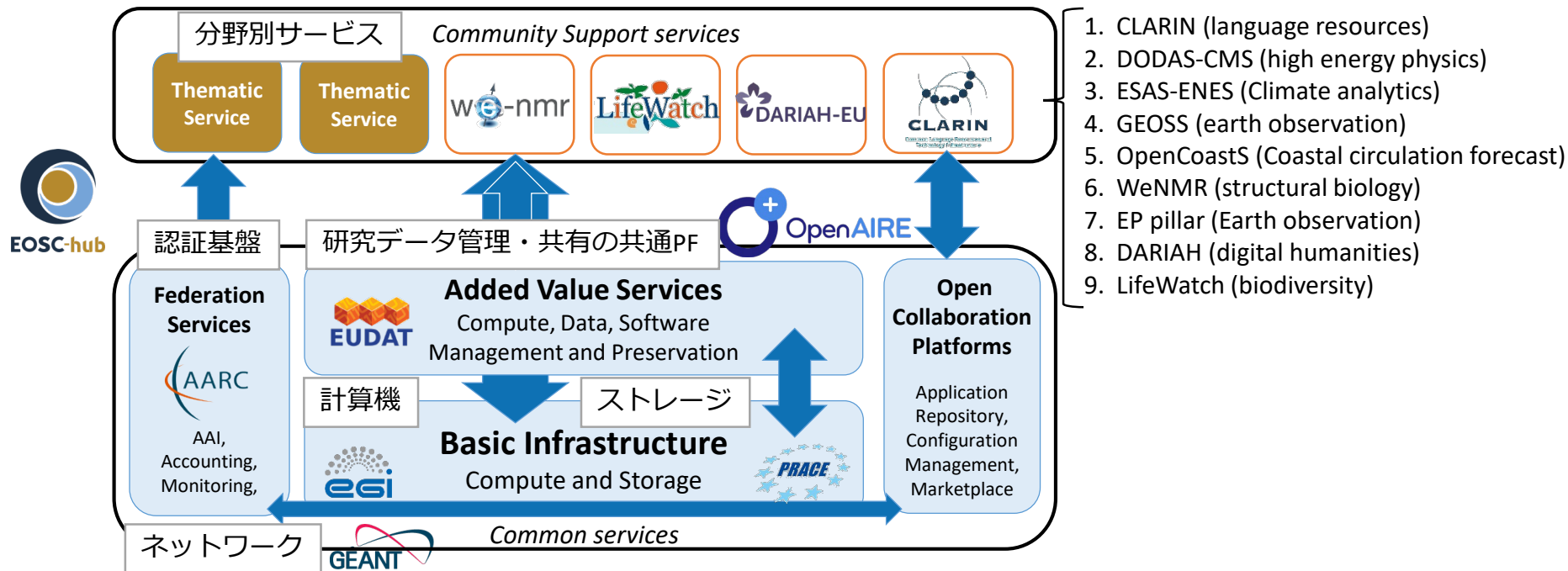


従来の文献を対象としたCiNiiとJAIRO Cloudを**研究データ**に対応

# European Open Science Cloud

- **従来**：個々の大学や研究機関でEインフラを構築
- **今後**：個々のインフラを繋ぎEU全体でEインフラを構築
  1. 既存のEインフラをベースにネットワークから分野別DBまでを有機的に結合
  2. 人社系へのサポートなどロングテールにおける活用も考慮
  3. 産業界との連携も考慮

## European Open Science Cloud (EOSC)



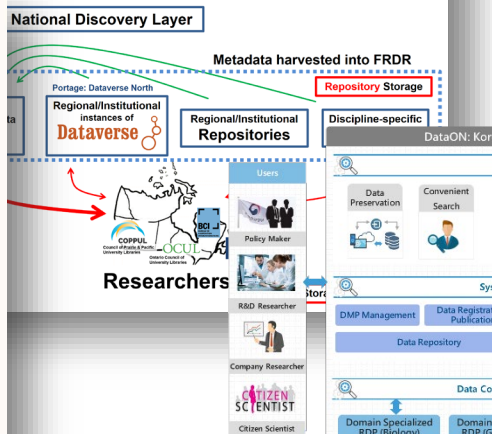
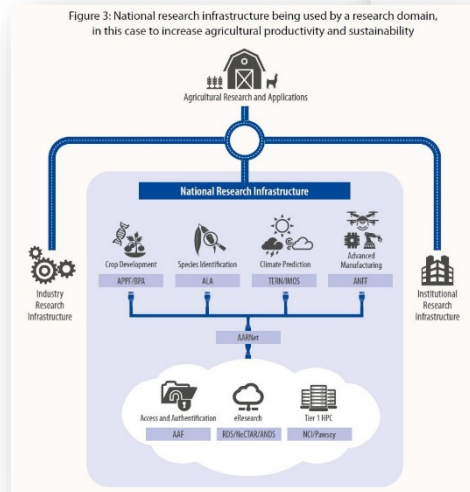
# 国レベルでの研究データ基盤開発の加速化

- Australian Research Data Commons
- Canadian Federated RDM
- Korea Research Data Platform
- Malaysian Open Science Platform
- African Open Science Platform

## African Open Science Platform (AOSP)

- Aims to be a coordinating initiative - interests, ideas, people, institutions and resources needed to advocate and advance open science in Africa
- Pilot Phase – (2016 – 2019)
  - Landscape Study, Database, Creating Awareness, Advocate
- 4 Focus areas + Research data management

## Federated RDM Architecture

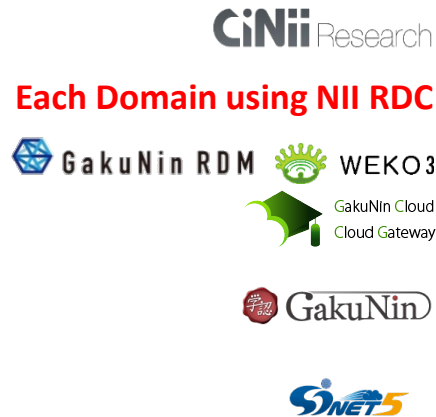


Capacity Building

# 研究データ基盤のアーキテクチャー

## • サービスレイヤ

このスタックがデファクト



共通ディスクバリサービス
分野別サービス
共通サービス
クラウド・HPC
VOプラットフォーム
認証フェデレーション
ネットワーク



## • 基盤整備のための連携

- サービス間連携：サービスレイヤ間、分野間
- 組織間連携：予算措置、共同ビジネスプラン

既存のEインフラを繋げて利便性や費用対効果を向上

これを政府レベルで積極的に推進

その**次**の方向性は？

# 欧州調査

---

日程：2019年6月17日～20日

メンバー：青木学聡（京大）、藤原一毅（NII）

## 1. スイス連邦工科大学チューリッヒ校 (ETH Zürich)

- 相手方: Caterina Barillari 氏 (Scientific IT Services) ら4名
- 研究データ管理システム OpenBIS の開発・運営に関する情報交換、および同校図書館が運営するRDMサポート業務に関する情報交換

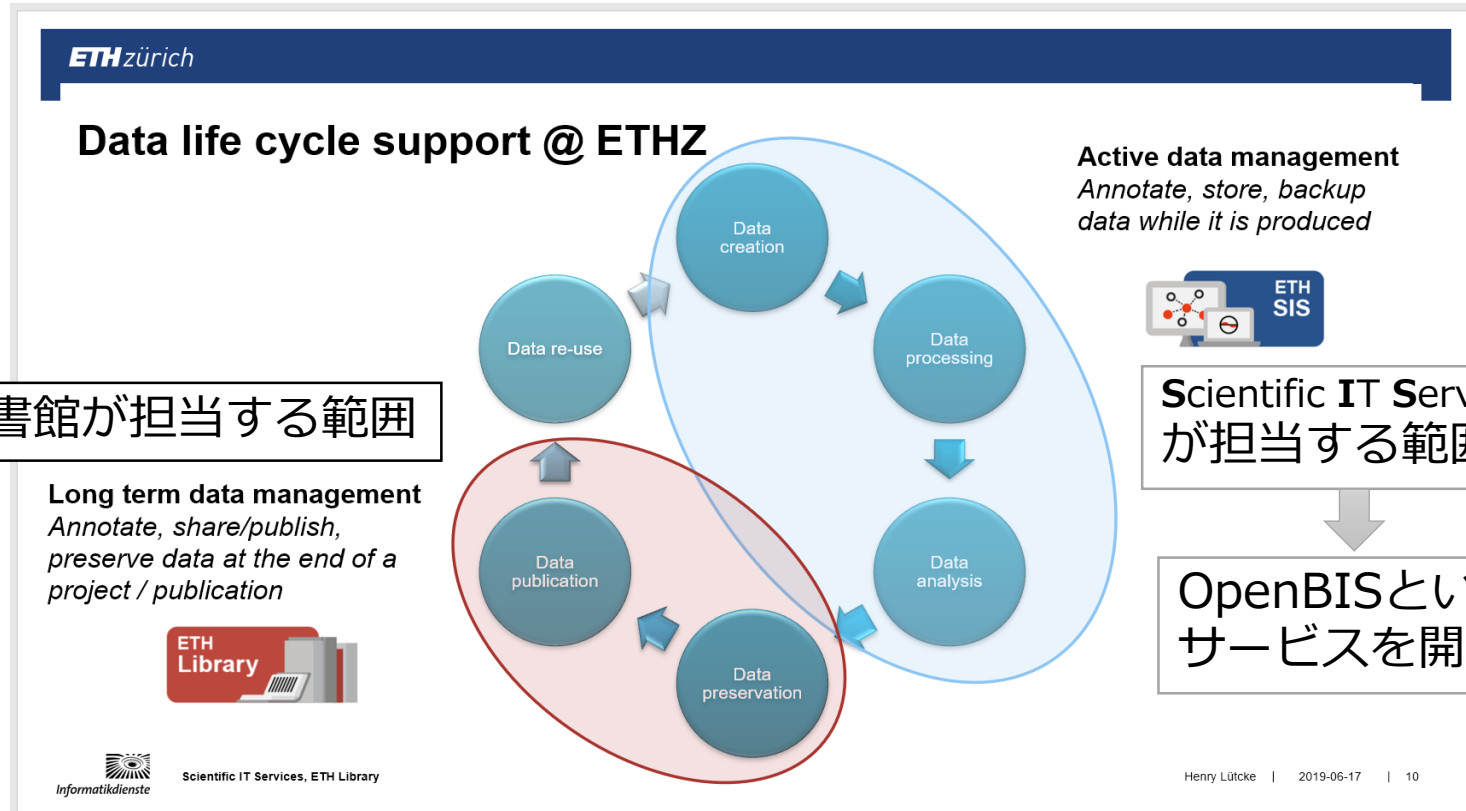
## 2. スイス連邦工科大学ローザンヌ校 (EPFL)

- 相手方: Giovanni Pizzi 氏 (Laboratory of theory and simulation of materials) ら3名
- 研究データ解析システム AiiDA の開発・運営に関する情報交換

## 3. ゲオルク・アウグスト大学ゲッティンゲン (GWDG)

- 相手方: Ramin Yahyapour 教授 (Managing Director) ら3名
- 同大学が運営している研究支援ICTサービスに関する情報交換

# ETHZ : 図書館と基盤センターによるサポート



LibraryとSISが合同で、教育プログラムも実施

- RDM Summer School (5日間の集中コースにより単位認定)
- RDM Workshop

個別のツール、システムの利用方法についても、SIS、Libraryがそれぞれセミナーを精力的に開催



# ETHZ : OpenBISの主な機能

---

研究室での統合的な情報管理を実現するために以下の機能を提供（2019年現在）

- インベントリ（ファイル、ドキュメント）サーバー
  - 研究室のメンバー間でのデータ共有
- **電子ラボノート**（Electronic Laboratory Note, ELN）
  - インベントリオブジェクトへのリンク埋め込みなど、ラボのソースと関係を明確に記述
  - オブジェクトブラウザにより、オブジェクト間の依存関係、ワークフローを可視化
- Jupyter Notebook等、**オンサイト解析環境**との連携
- **外部ワークフローエンジン**との連携
- 外部の大規模データストレージへの参照機能

# 豪州調査

---

日程：2019年6月17日～20日

メンバー：込山悠介（NII）、常川真央（NII）

## 1. [The University of Queensland \(UQ\)](#)

- 相手方: Joe Shapter教授(Pro Vice Chancellor Research Infrastructure)ら4名
- 研究データ管理システム UQ RDMの開発・運営に関する情報交換

## 2. [University of New South Wales \(UNSW\)](#)

- 相手方: Maude Frances氏(Associate Director, Scholarly Communications and Repositories)ら4名
- 同大学における研究データポリシー、研究データ管理システム、RDMトレーニング教材作成の取り組みに関する情報交換

## 3. [University of Sydney \(USYD\)](#)

- 相手方: Andrew Janke氏(Associate Director, Research Technology)ら4名
- 同大学が提供するデータサイエンスのトレーニングプログラム Sydney Informatics Hub の運営に関する情報交換

# UNSW : ResData

FOR MORE INFORMATION GO TO <https://research.unsw.edu.au/research-data-management-UNSW>

### Storing your Research Data

Storage Platforms	Key SUPPORTED				Currently UNSUPPORTED		
	UNSW OneDrive & Teams	UNSW eNotebook	Data Archive	Home Drive Shared Drive	Cloudstor	Dropbox	Local Storage <sup>1</sup> (eg. external/USB drive)
Storage Type	Day-to-Day	Day-to-Day	Long-Term	Day-to-Day	Day-to-Day	Day-to-Day	Day-to-Day
Suitable Data Classification	Public	Private	Private	Private	Public	Private	Private
Stored in Australia	✓	✓	✓	✓	✓	✗	?
Backup & Disaster Recovery	✓	✓	✓	✓	✗	✗	✗
Synchronising with Local Copy	✓	Not Applicable	Not Applicable	Not Applicable	✓	✗	Not Applicable
External Collaborator Access	✓	✓	✗	✗	✓	✓	✗
Storage Limit	5 TB/User	Unlimited	Unlimited	Unlimited	1TB/User	\$\$\$	?
Version Control	✓	✓	✓	✓	✓	✓	✗
Recovery from Deletion	60 Days	No Data Deletion	No Data Deletion	10 days	✗	\$\$\$	✗
Past-Project Data Retention	> 7 years	Indefinitely	Indefinitely	> 7 years	?	\$\$\$	?

\* Local devices vary greatly in their configuration and security. Contact [info@unsw.edu.au](mailto:info@unsw.edu.au) to find out which data classification is suitable for your local device (eg. laptop) (if relevant)

Legend: ? Unknown or Data Dependent, ✗ Highly Sensitive Data, ⚠ Sensitive Data, 🛡 Private Data, 🌐 Public Data

<sup>1</sup> For Sensitive and Highly Sensitive data, data encryption and/or other storage must be reviewed. Please refer to the UNSW Data Handling Guidelines for more information. If you have any Highly Sensitive data, or Research Data Management enquiries, please contact [rdm@unsw.edu.au](mailto:rdm@unsw.edu.au).



### ResData

Welcome Elizabeth Stokes (Library Repository Services, University Library)

- Managing data
- Publishing data
- Publish dataset record
- Search datasets
- Log out

#### Datasets

Recent records (maximum of 5 shown)

Title	Updated	Status	Action
Taxonomy of enterological disease	20-01-2016	Pending	Manage
Muscle protein synthesis and nutrition	29-02-2016	Pending (DOI)	Manage
Assessing micronutrient status in the presence of inflammation	29-02-2016	Draft	Manage
Micronutrient deficiency in adolescents	29-02-2016	Draft	Manage

#### Projects/grant(s)

You are participating in the following project/grant(s):

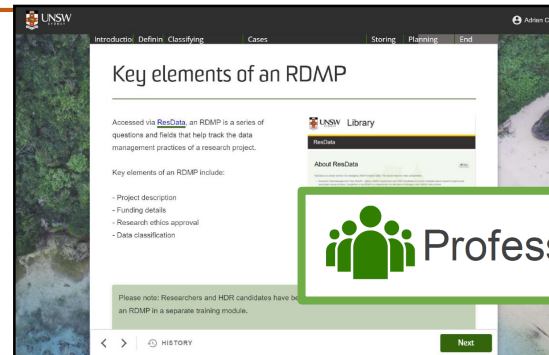
Name	Funding Scheme	Role	Action
Role of micronutrients in treating inflammatory bowel disease	ARC	Contributor	Manage

- 判別表 “Storing your Research Data” を用いて研究者自身がデータ性質を格付け
- データの種類ごとに適切なストレージをシステムが提供
- データアーカイブシステムと連携したDMPシステムを提供研究後のデータアーカイブ領域の自動確保
  - ResDataで作成したDMP情報に基づいてUNSW運用のストレージ「DataArchive」をリクエスト可能
  - 主に研究後の研究データ保管が円滑に行われるよう設計

# UNSW : RDMoT



Research Data Management online Training



- RDMオンライン教材を提供
- LMSを導入して独自にオンライン教材を作成
- RDMの一般知識と併せて学内ツールのインストラクションも実施
- 学習履歴の分析：離脱率や回答の流れを集計・可視化
- Hacky Hour：RDMを含めた研究ツールのレクチャーとワークショップ
- ResBaz：国際的なHacky Hourとしてのカンファレンス（研究者、大学院生のマッチングメイキングの役割もある）

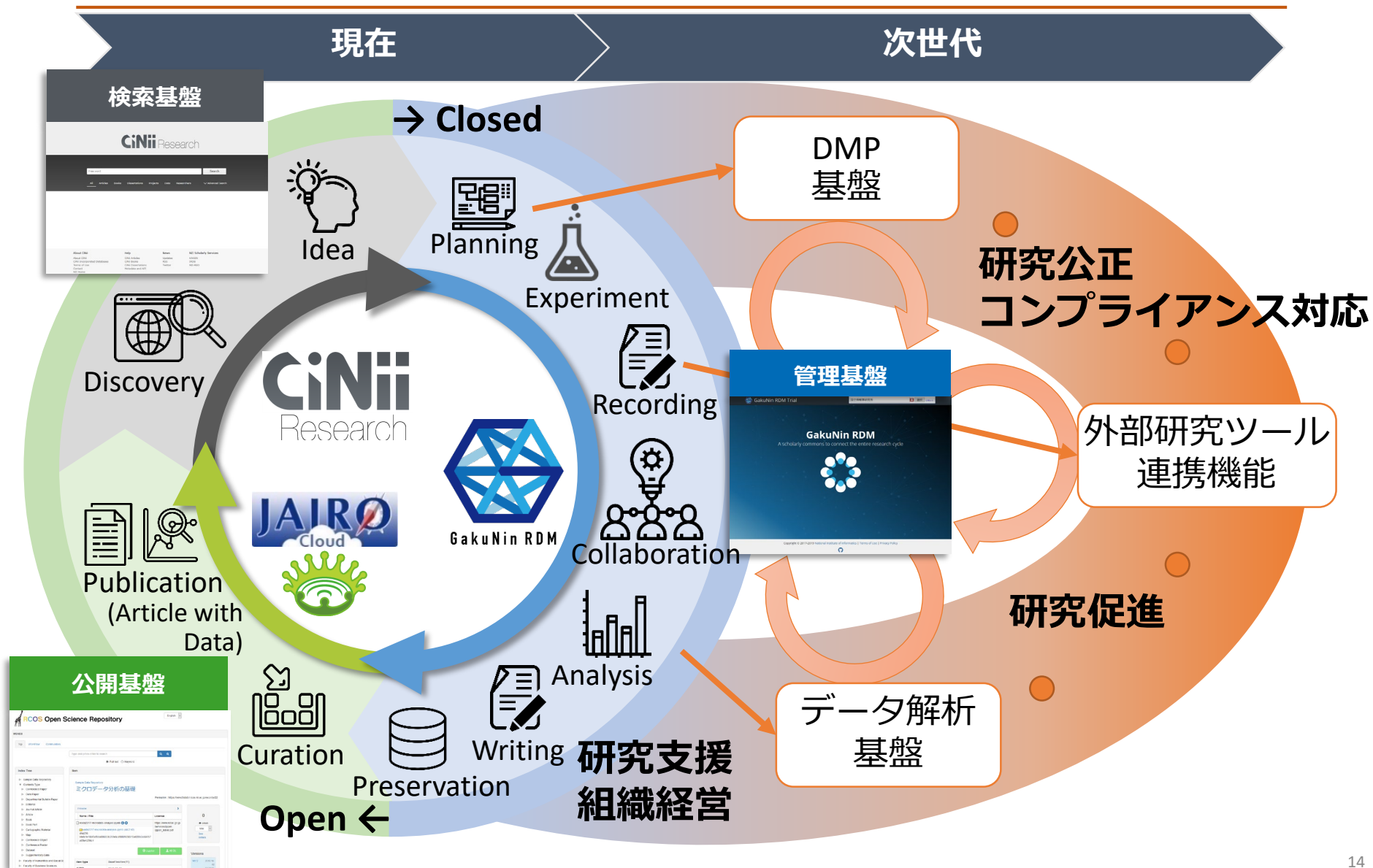
# 海外における先進的な取り組み

- ラボやプロジェクトでの研究データの適切な共有環境を組織として提供 ⇨  GakuNin RDM

+

- ① 研究を支援するためにデータ解析システムや電子ラボノートといった研究に更に密接なサービスを提供
- ② 研究公正やコンプライアンスに対応するために研究データ管理を機械的にサポートする基盤
- ③ 人材を育成するためにトレーニングやキュレーション支援の充実化

# 新しい研究データ基盤

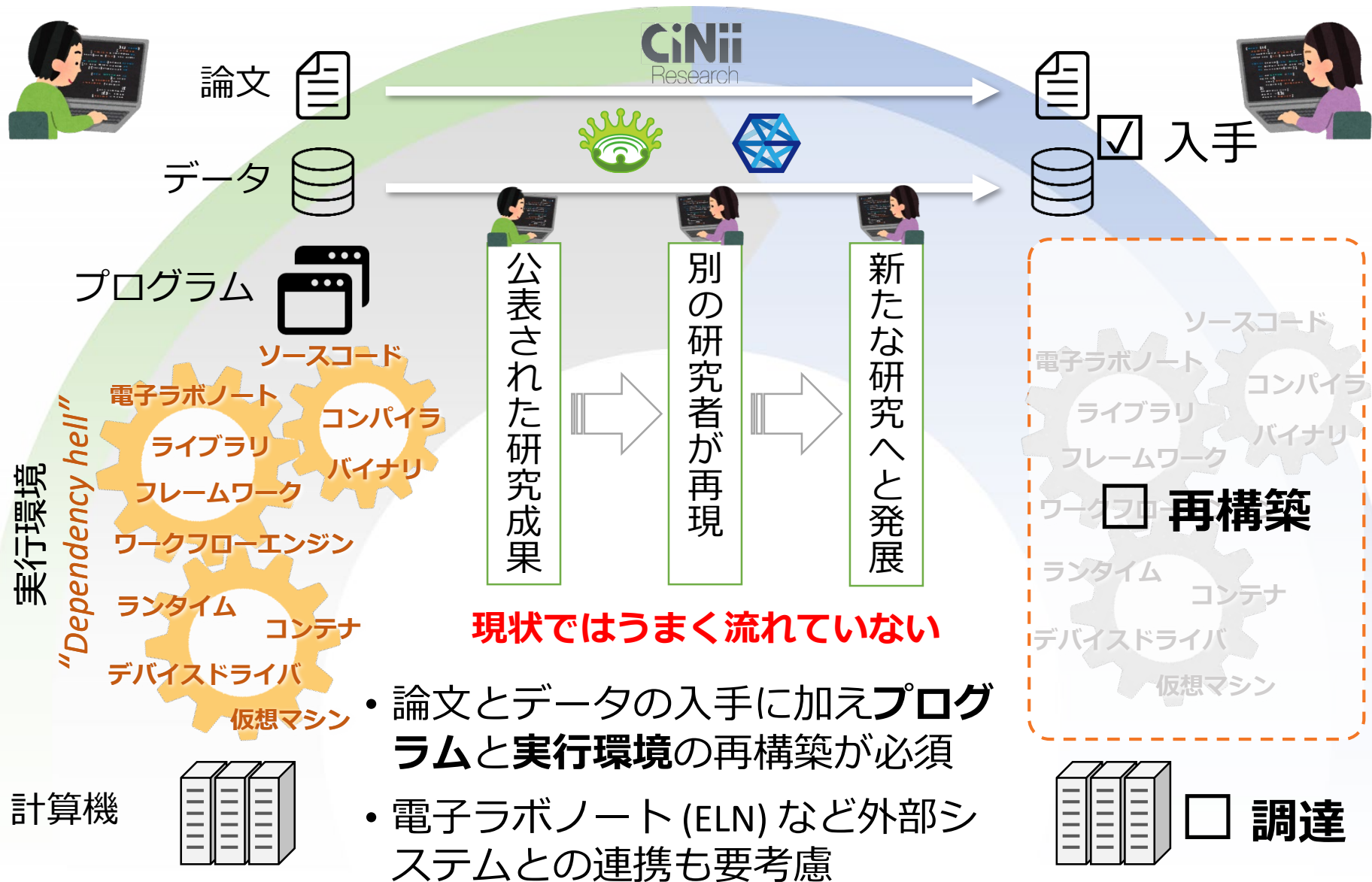


# 研究データ基盤整備の方向性

---

1. 研究データ基盤の高度化
  1. データ駆動型研究促進
  2. コンプライアンス対応
  3. 研究支援・組織経営
2. 研究データ管理・公開支援者の育成
  1. データキュレーション基盤の構築
  2. 人材育成のための教材・教育基盤の構築

# 1-1 研究促進のための高度化 ～オープンサイエンスの環をつなげる～





# 1-1 データ分析環境の自動再構築

～研究のボタンタッチを簡単・確実に～



データ分析に必要な環境を、コンテナ化技術を用いて簡単に自動構築可能

GakuNin RDMから必要な情報をパッケージ化して、WEKO3から簡単に公開可能

自動環境再現システムにより、異分野の研究者でも簡単に再利用可能

# 基盤構築から見た1-1と1-2の関係

---

データ駆動型研究促進



研究成果のパッケージ化



コンプライアンス対応

# 1-2 コンプライアンス対応のための高度化 ～ DMPと研究の実態が同期するエコシステムの構築～



DMPに基づく研究環境の整備

再現性情報をDMPに関連付け



DMPを研究中に随時参照し見直し

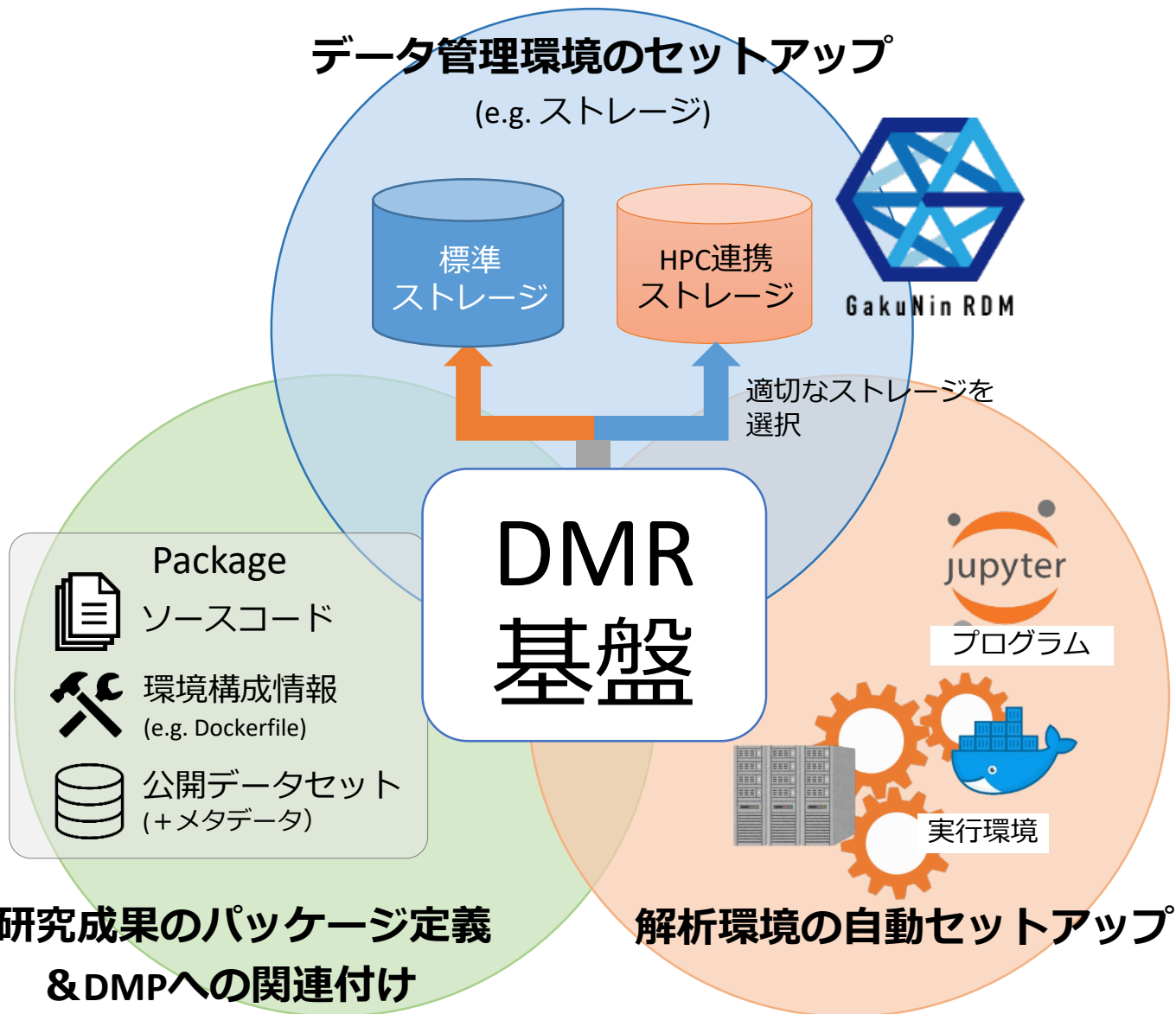


## Progressive DMP

研究成果のパッケージ定義&DMPへの関連付け

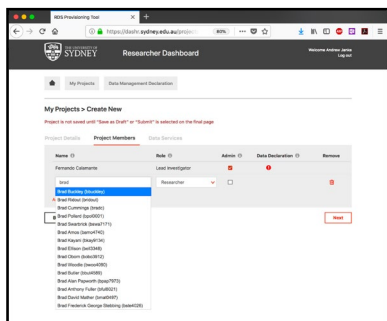
# 1-2 コンプライアンス対応のための高度化

～実現のための機能例～



# 1-3 海外における組織経営(IR)への取り組み

- 豪州大学におけるデータ管理基盤



- 学内に蓄積されたDMPを分析
- 学内の研究活動を可視化
  - 例1: 学部間の共同研究の活発度
  - 例2: 学部ごとのストレージ使用量
- 分析結果をもとに経営戦略に活用
  - 例1: ストレージ予算の説明など

- 海外ディスカバリーサービス



- OpenAIRE Research Community Dashboard
  - 特定の研究コミュニティに関連する OpenAIRE Research Graphや統計情報を提供するゲートウェイを構築
- CHORUS Institution Dashboard
  - 助成機関等にオープンアクセスコンプライアンスを追跡可能とさせる。JSTを含め12機関が参画

# 1-3 研究支援・組織経営のための高度化

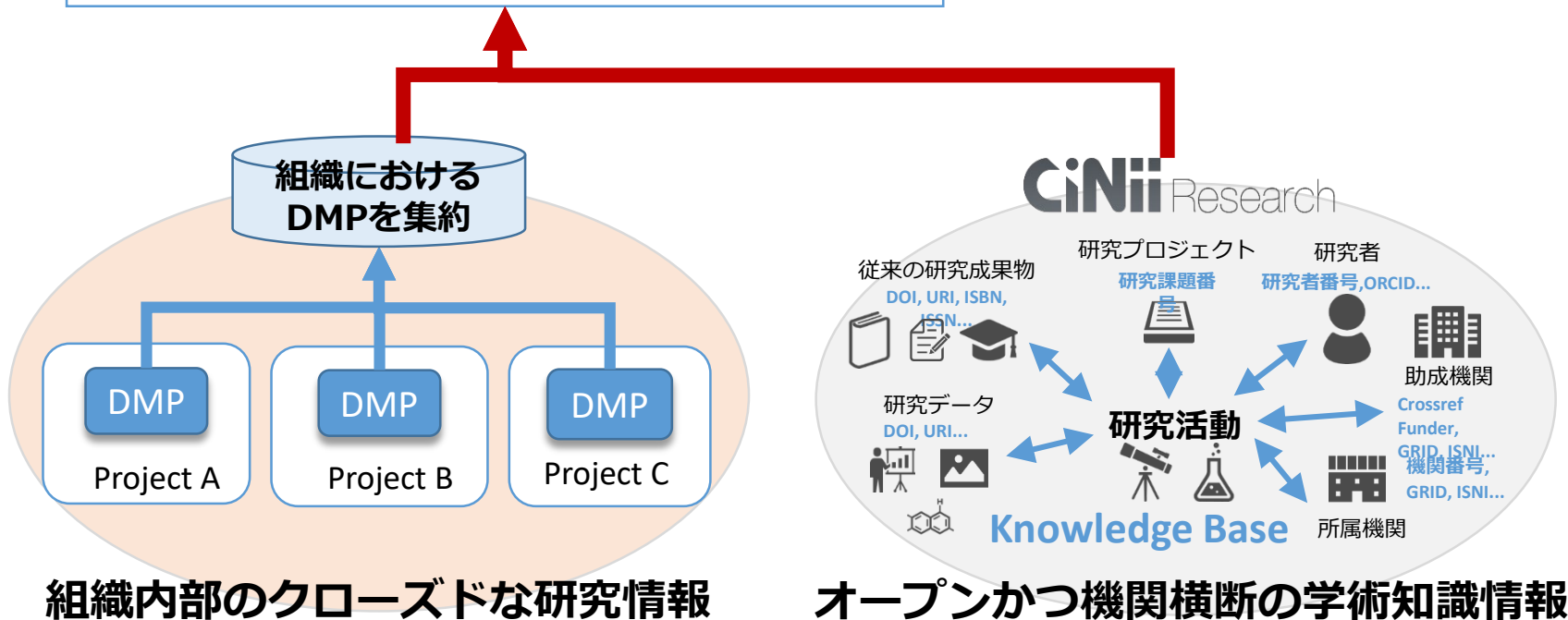
個々の研究者をサポートするだけでなく、研究者間の関係性構築や研究機関の効果的な運営にも生かされる機能提供

## 研究データ・ダッシュボード機能の提供



## 研究機関の戦略立案支援

- ・ データのインパクトを評価
- ・ 共同研究の可視化
- ・ データ人材の発掘



組織内部のクローズドな研究情報

オープンかつ機関横断の学術知識情報

# 研究データ基盤整備の方向性

---

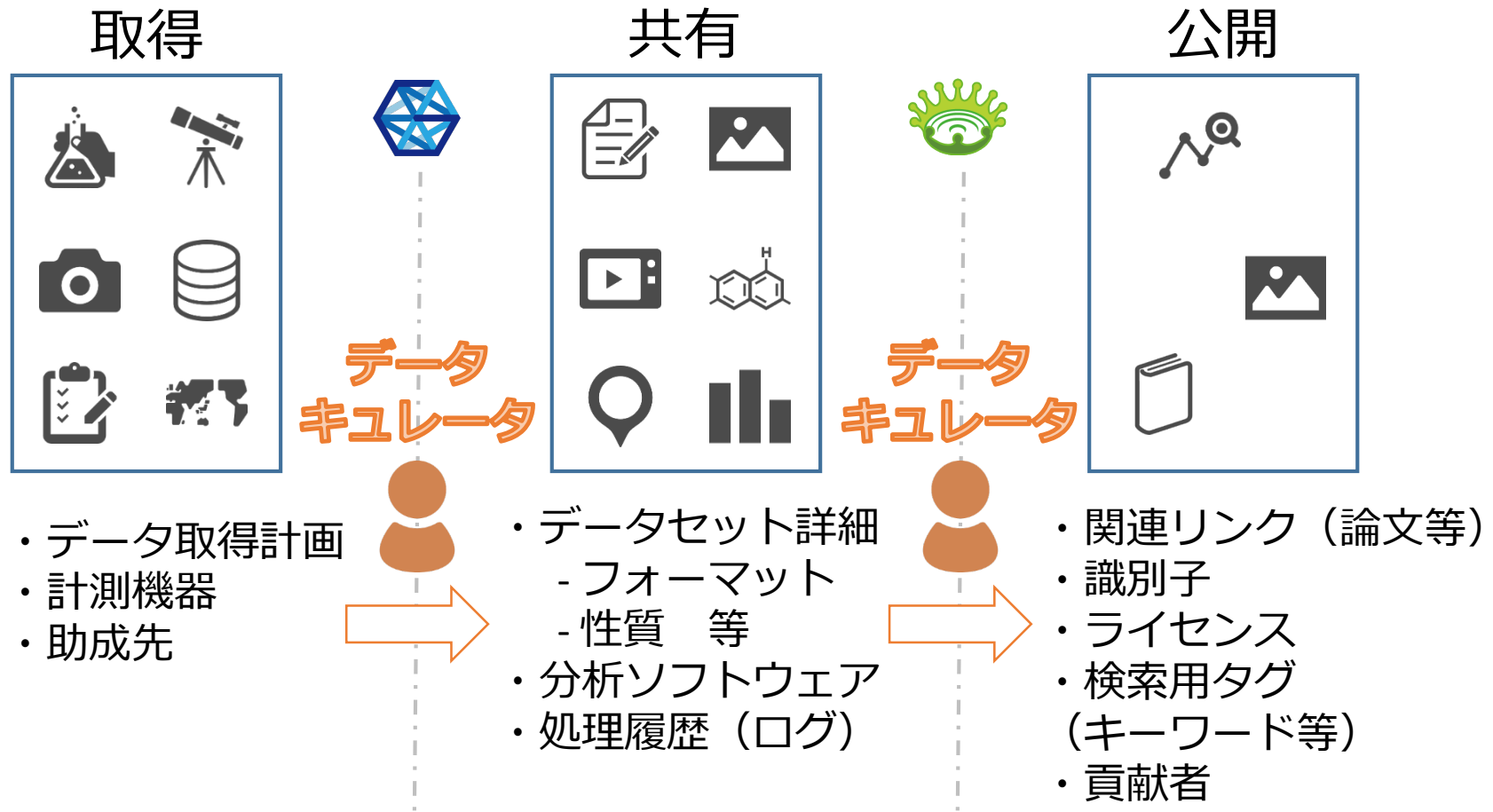
## 1. 研究データ基盤の高度化

1. データ駆動型研究促進
2. コンプライアンス対応
3. 研究支援・組織経営

## 2. 研究データ管理・公開支援者の育成

1. データキュレーション基盤の構築
2. 人材育成のための教材・教育基盤の構築

# 2-1 データキュレーションの流れ



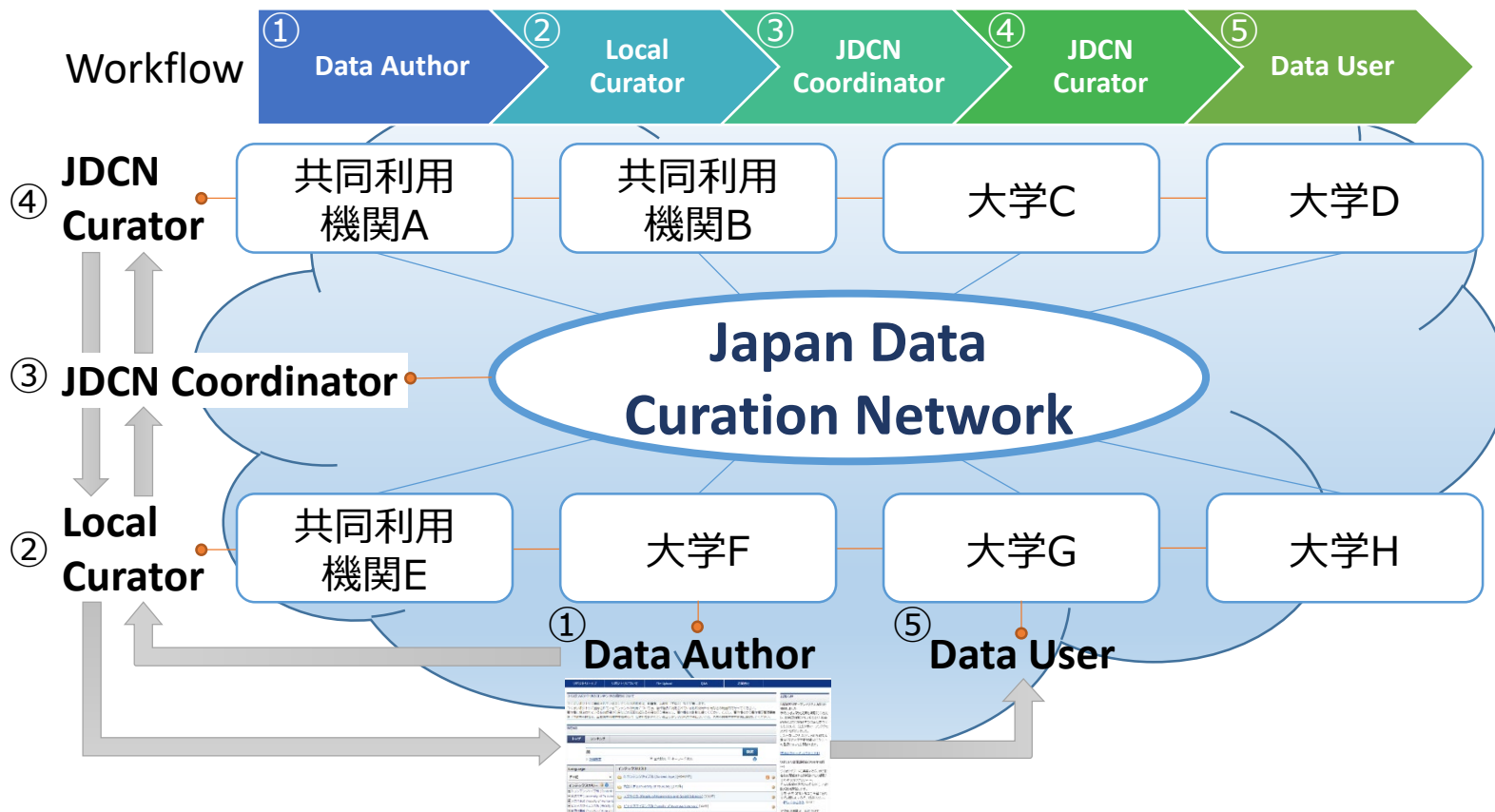
多様なデータをキュレーションする人材発掘・育成が課題



# 2-1 データキュレーションの実践

## ～人的ネットワーク基盤構築～

- 大学図書館員、研究支援職員、研究者（データキュレーター）を含む人的ネットワークを構成する基盤
- 単一の機関でカバーしきれない多様な専門分野のデータキュレーションを、大学共同利用機関を含む複数機関間で相互に支援
- 米国の複数大学図書館で実践されているほか、オランダ・カナダでも試行中



## 2-1 データキュレーション基盤の機能

- 基盤は5つの主機能から構成

ワークフロー機能

### 人材プール機能

- 一定の認証を経たキュレーターをプール
- シニア - ジュニアへの知識継承の場を提供

### マッチメイク機能

- キュレーター情報・活動ログから最適なキュレーターを選択
- 参加館のリソース配分最適化

### 活動状況レポート機能

- キュレーターの貢献度を数値化し、参加機関へフィードバック

### 認証マーク発行機能

- データの認証マークを発行し、キュレーターID/作業履歴のメタデータを併せて外部システムへ引き渡し

- 人的ネットワークの中核にはデータコーディネータを配備
- 人材育成のために教育・研修システムと連携

# 2-2 人材育成のための教材開発と利活用促進



教育コンテンツの**拡充**

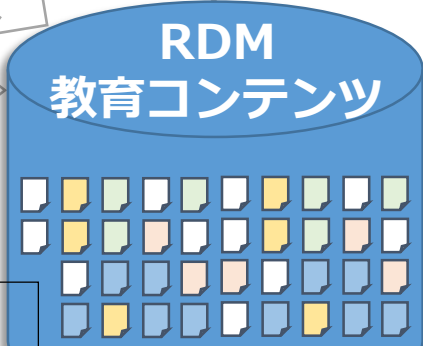
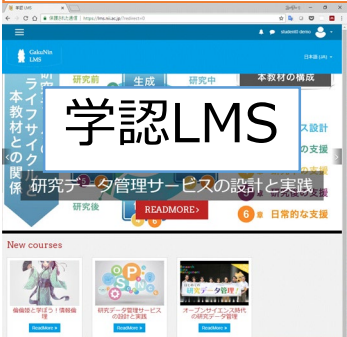
再利用カスタマイズ可能な  
**マイクロコンテンツ**教材化

RDMスキル/支援スキル  
**コンピテンシー**の策定

- DONE** (基礎編)  
「OS時代の研究データ管理」
- DONE** (支援者向け)  
「RDMサービスの設計と実践」
- NEW** 研究者向け教材
- NEW** NII-RDC利用者向け実務教材



項目	内容
基礎編	OS時代の研究データ管理
支援者向け	RDMサービスの設計と実践
研究者向け	研究者向け教材
NII-RDC利用者向け	NII-RDC利用者向け実務教材



**NII** 大学共同利用機関法人 情報・システム研究機構  
**国立情報学研究所**  
National Institute of Informatics

コミュニティと連携・協力

学術情報NW運営・連携本部  
OS研究データ作業部会

J P C O A R  
オープンアクセスリポジトリ推進協会

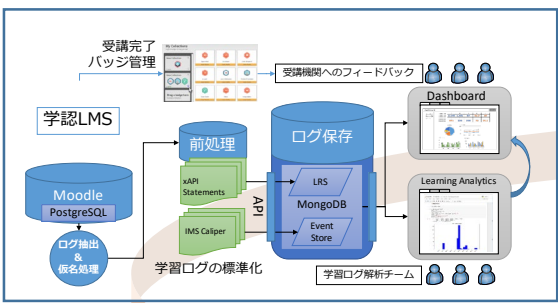
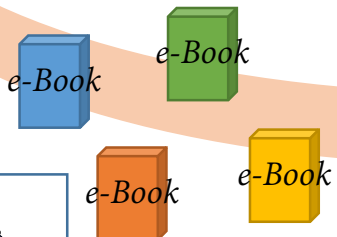
**AXIES** 大学ICT推進協議会

# 2-2 人材育成のための教育基盤



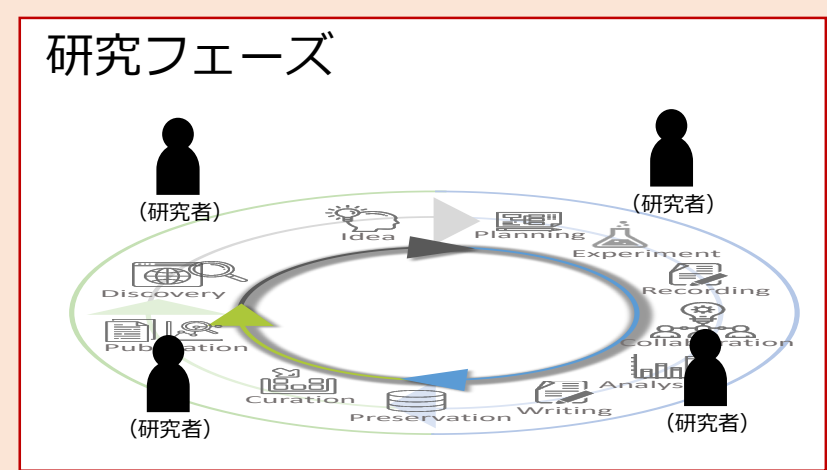
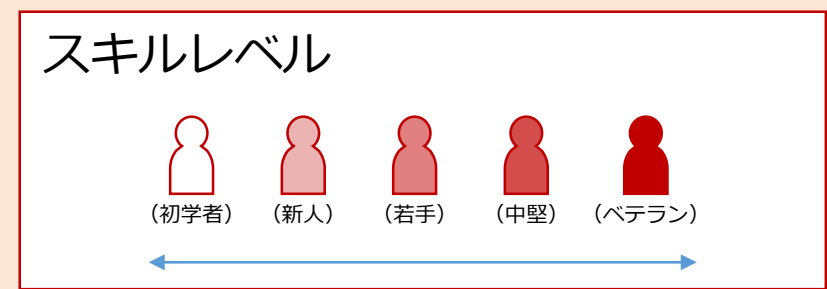
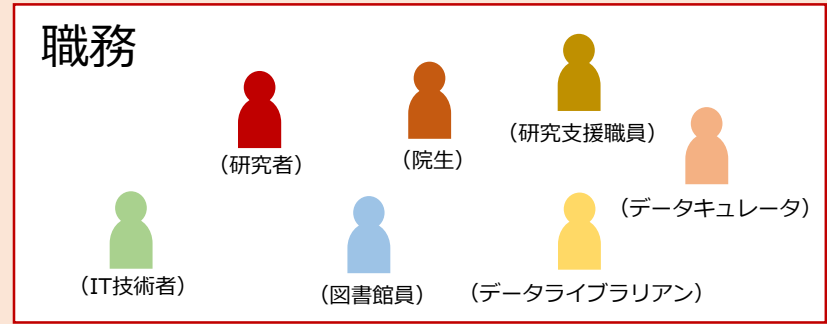
学習者に最適化した  
マイクロコンテンツ教材  
カスタマイズ機能

学習者の属性・状況に応じた  
教材リコメンド機能



## ラーニングアナリティクス機能も配備

- ・ 学習状況の解析に基づく教材改善、学習支援
- ・ 学習修了認定バッジ発行



# まとめ

---

- 海外における研究データ基盤の整備の動向
- NII RDCの高度化の方向性
  - 研究データ基盤の高度化
    - データ駆動型研究促進
    - コンプライアンス対応
    - 研究支援・組織経営
  - 研究データ管理・公開支援者の育成
    - データキュレーション基盤
    - 教育基盤

**RCOS**  
yamaji@nii.ac.jp