

データマネジメントプラン(DMP)駆動型の データ環境構築サービス

2020年06月09日

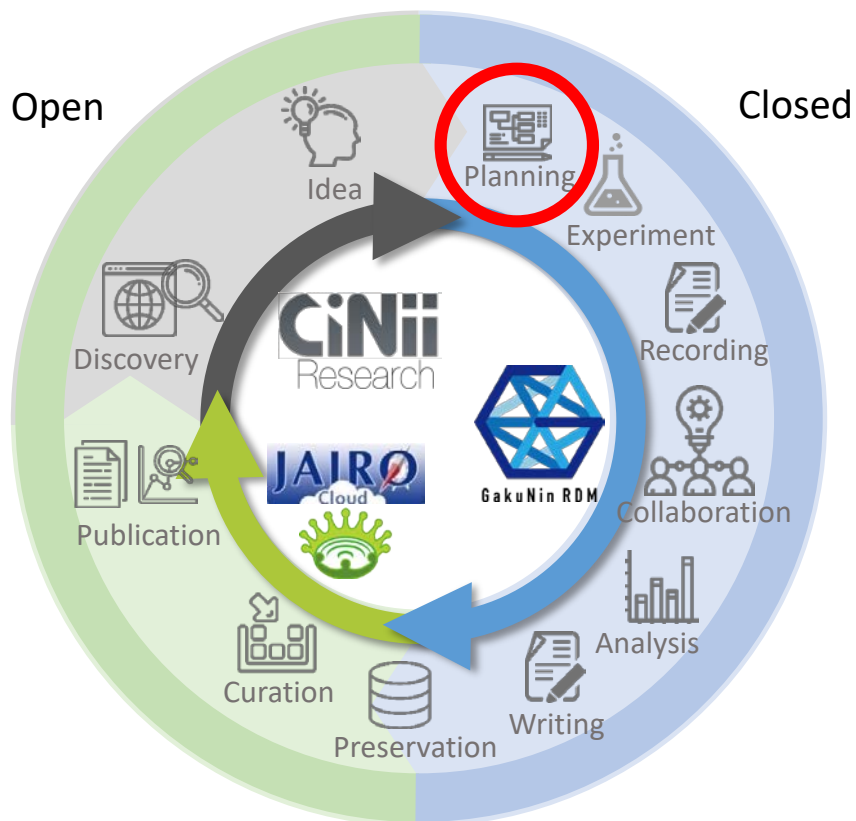
国立情報学研究所

オープンサイエンス基盤研究センター

常川真央

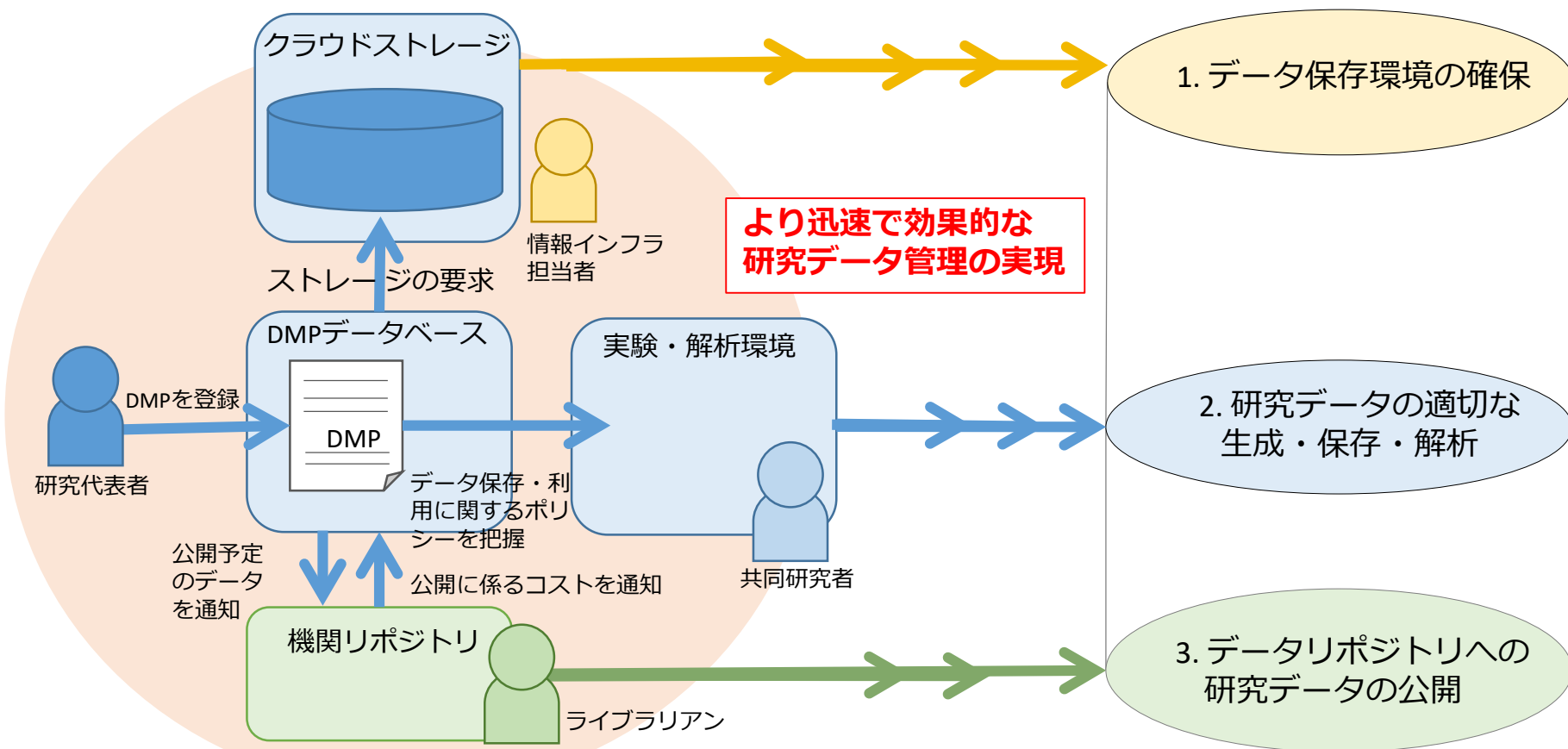
背景: データマネジメントプラン(DMP)

研究プロジェクト中およびプロジェクト完了後のデータの処理方法を概説する正式な文書 [EC, 2018]



- **研究データのライフサイクルの第1歩にあり、研究データ管理に不可欠な要素**
- DMPをめぐる動向
 - 国際動向：データマネジメントプラン(DMP)が研究データ管理の中核的な情報源として重視
- 国内ではJST、AMED、NEDOに加えJSPSもDMPの提出を義務化

国際動向：DMP活用によるデータ管理自動化の取り組み



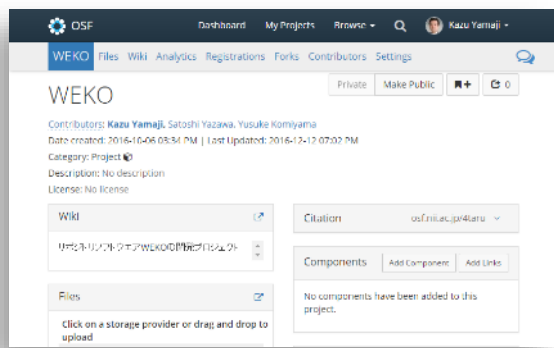
- DMP活用の取り組みとして、DMPをシステム同士が直接やり取りし、研究データ管理に必要な環境の用意を半自動化するソリューション(= Machine Actionable DMP)が提案
⇒ 現在、研究データ同盟(RDA)など国際的な組織が共通規格の策定やシステム開発を推進

現状の研究データ管理基盤

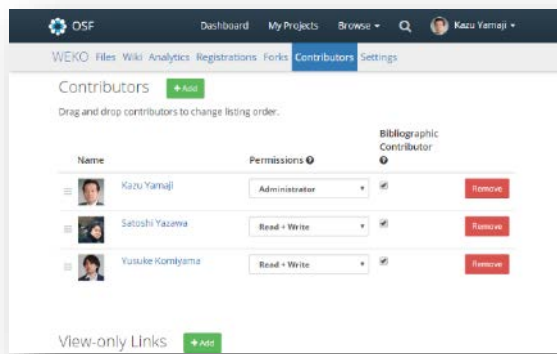


- 研究データ管理に必要なデータの保存・共有機能を提供
- **ただし、現状ではDMPに従った研究データ環境の整備や確認は手動で行う必要**

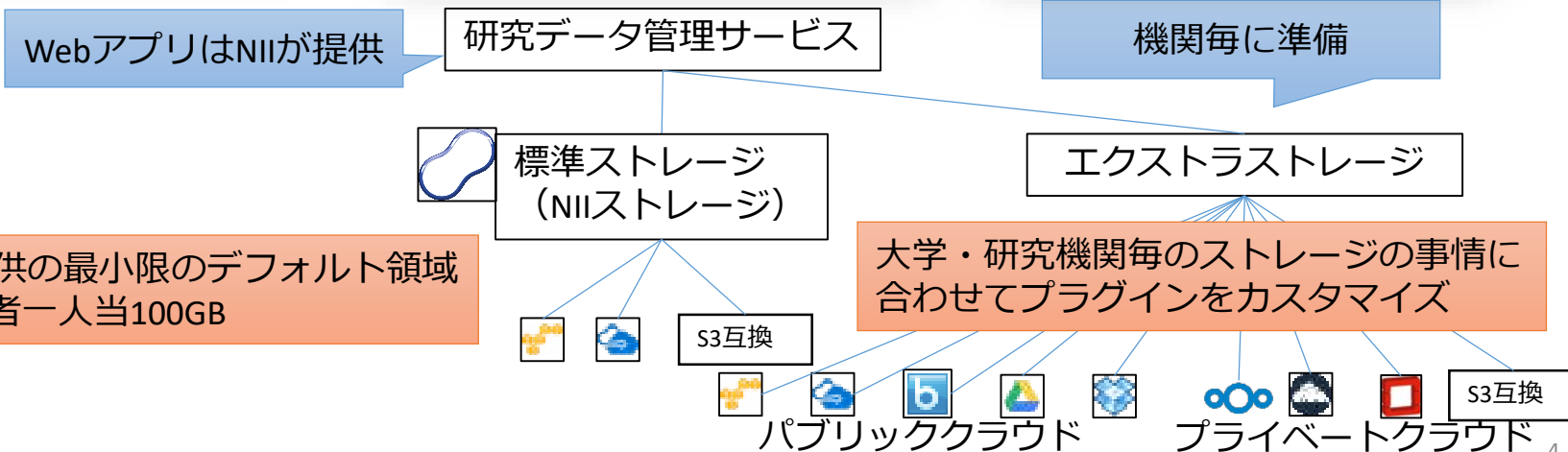
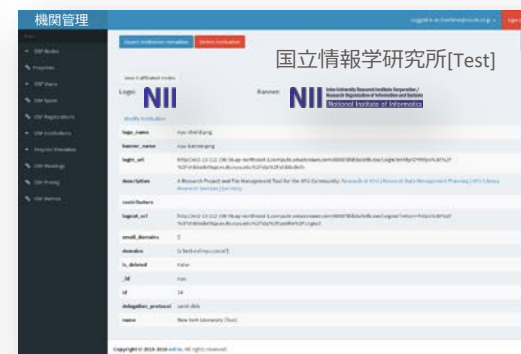
研究プロジェクト単位で
ファイルなどを管理



学認と連携しVO(仮想組織)
メンバーでファイルを共有

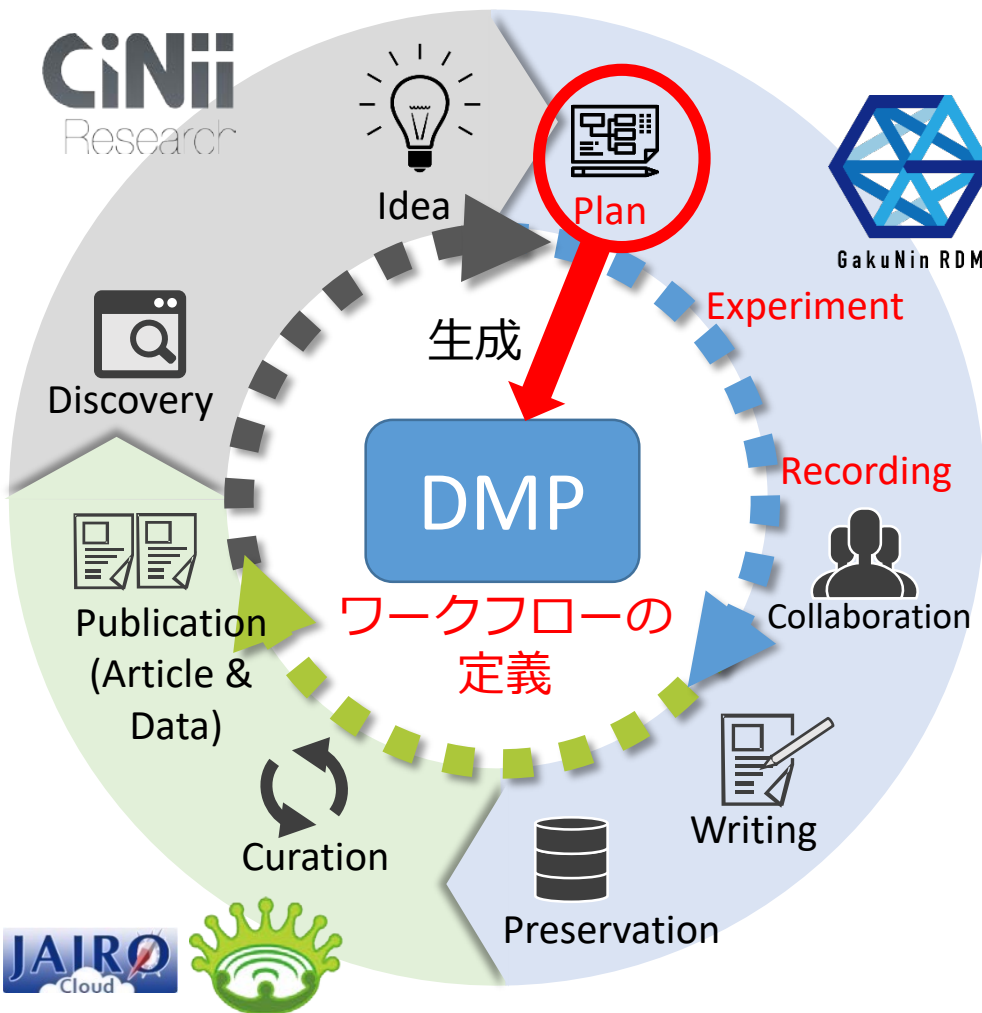


機関提供のストレージを利用し
研究証跡を保存・保護



NII Research Data CloudのDMP対応に向けたビジョン

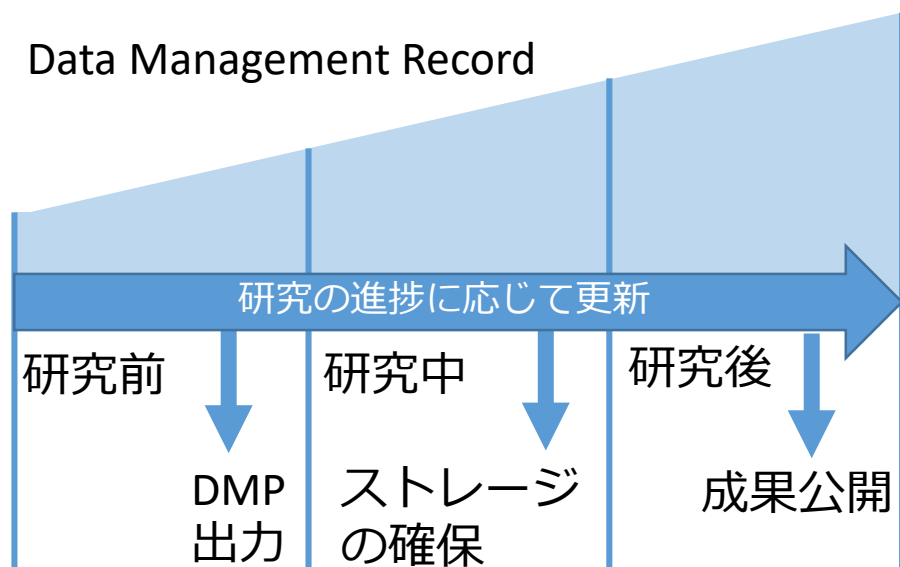
研究を着想したらすぐにデータライフサイクルの道筋を形作れる



- DMPを作成することで、NII 各基盤を活かした研究活動を展開可能にする
- ビジョン達成のためにNII RDCはどのような機能を備える必要があるか？
- 2019年度: 機能要件の検討にあたり海外DMP活用動向を調査

海外事例(1) : DMPの内部データモデル

Data Management Record (DMR) と呼ぶコンセプトを採用し、研究の全過程におけるDMPの漸進的な更新と活用を実現



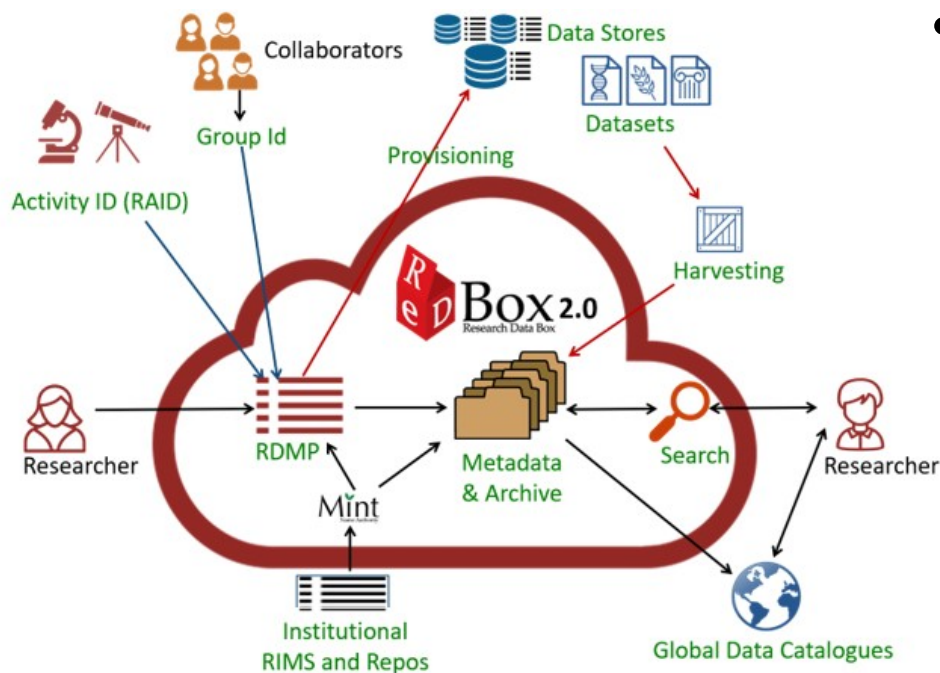
- 研究プロジェクトのRDM状況を「レコード」として記録
- 研究の進捗に応じて更新し、情報を連携サービスに受け渡し
- DMPの更新履歴が保存され、研究データ管理の記録として利用可

USYDによる応用事例:

DMRの情報を集計し、機関のIRに活用（プロジェクトごとのストレージ使用状況、学部間の共同研究の状況、研究データの公開状況）

海外事例(2): システムアーキテクチャ

DMP管理システムを中心として研究支援サービスを制御するアーキテクチャを採用

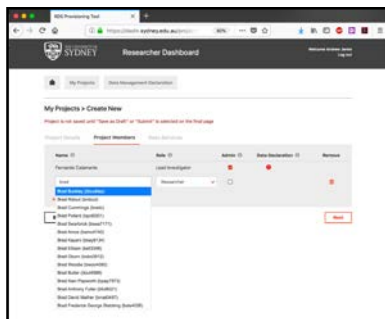


- シドニー工科大学によるRDMシステム “Stash”
 - ベースとしてReDBoXを導入
 - DMPを作成・管理
 - DMPに基づいて外部サービスに対して研究データ環境をプロビジョニング
 - 連携先で生成されたデータを抽出し、メタデータを作成

ReDBoX 2.0: 研究データの計画, 管理, 公開を支援することを目的としたオープンソースソフトウェア. Provisioner APIという共通APIを介して多様なサービスと連携できるように設計されている [QCIF, 2020]

海外事例(3) : 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

DMP活用の目的

DMR活用事例から以下の3点の目的を抽出

1. **コンプライアンス**： 研究データ管理計画を説明する責任を果たすための文書として出力し， 研究助成機関へ提出する
2. **研究促進**： RDMの現在の状況を組織内に伝達し， 最適な研究データ管理環境の構築・連携を促進する
3. **組織経営**： 学術研究機関のIRに関する情報源としてDMPを使用し， ストレージのコスト計算や研究成果情報の集約を支援する

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



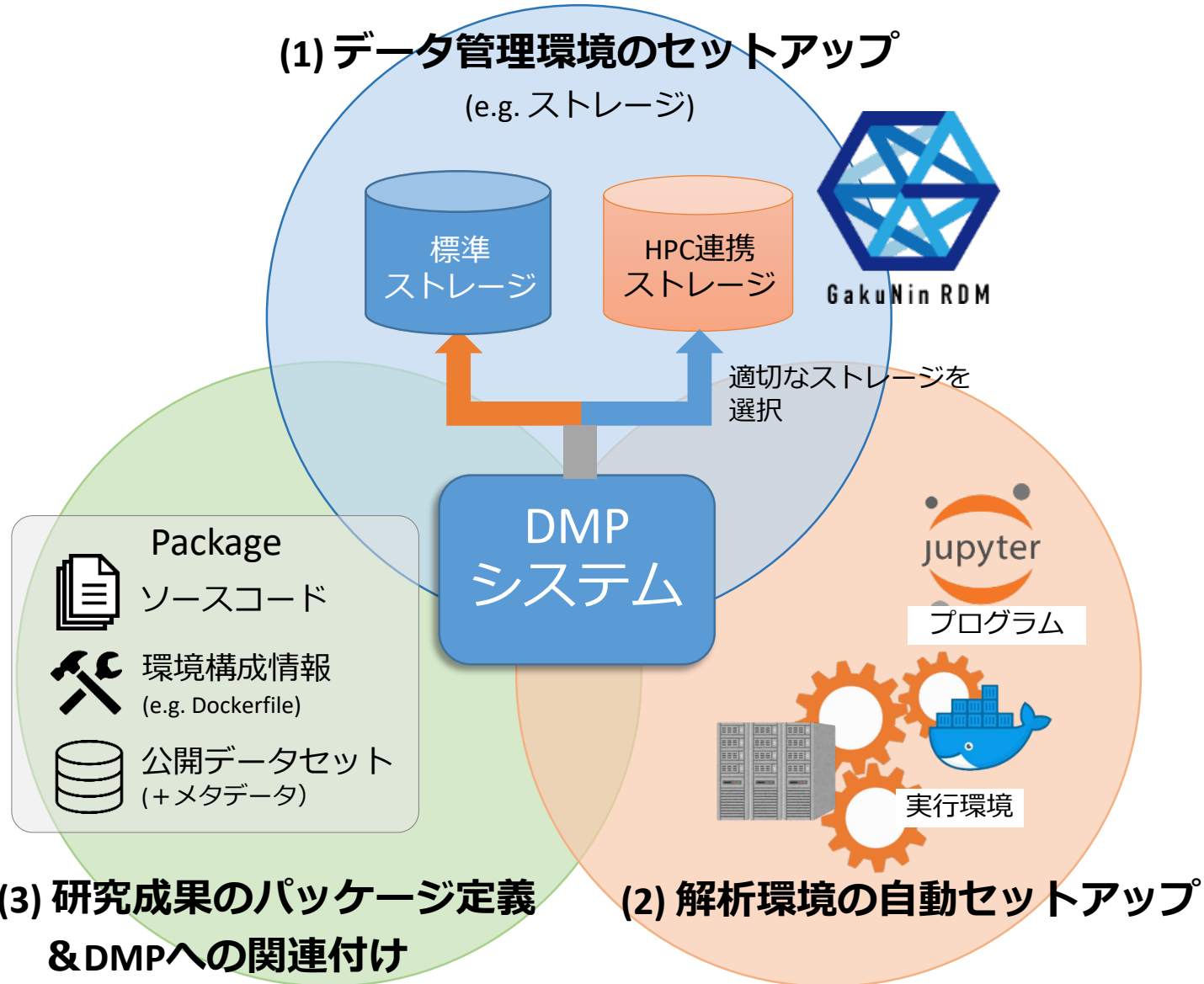
DMPを研究中に随時参照し見直し



Progressive DMP

(3) 研究成果のパッケージ定義DMPへの関連付け

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～



研究支援・組織経営のための高度化

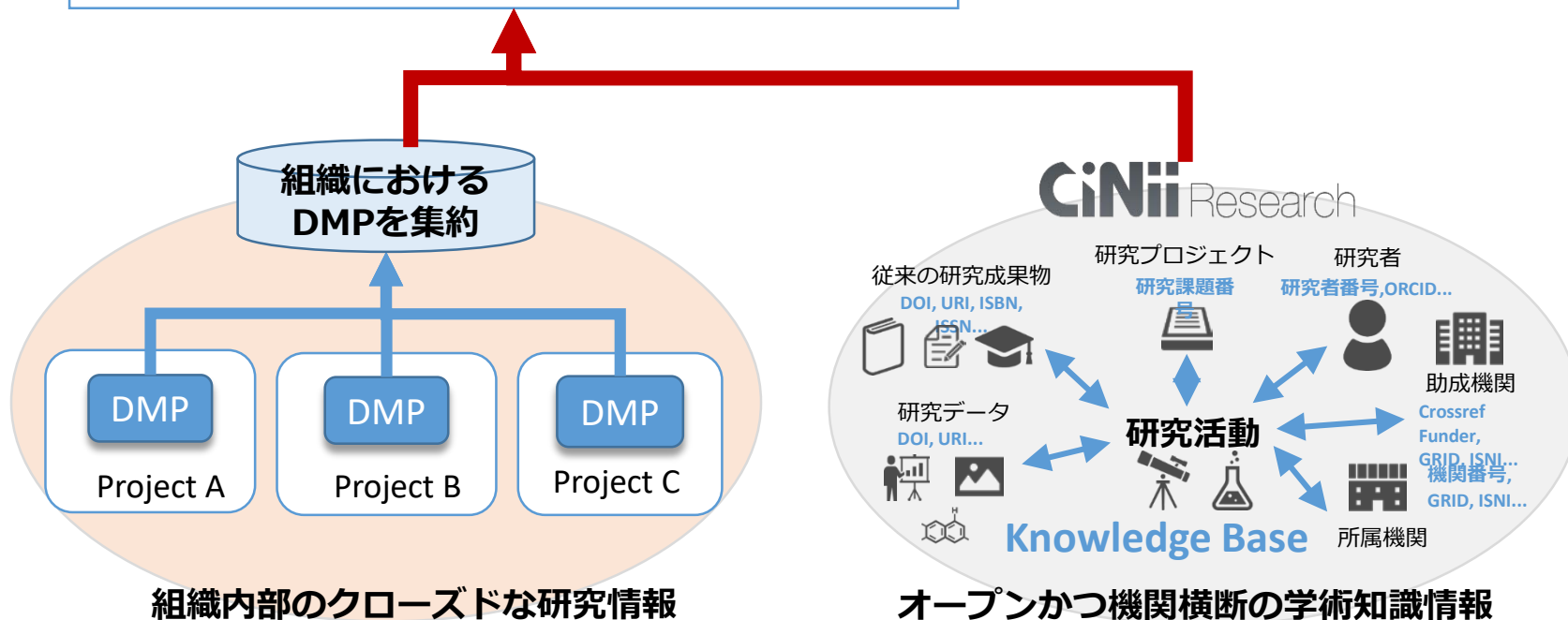
個々の研究者をサポートするだけでなく、研究者間の関係性構築や研究機関の効果的な運営にも生かされる機能提供

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

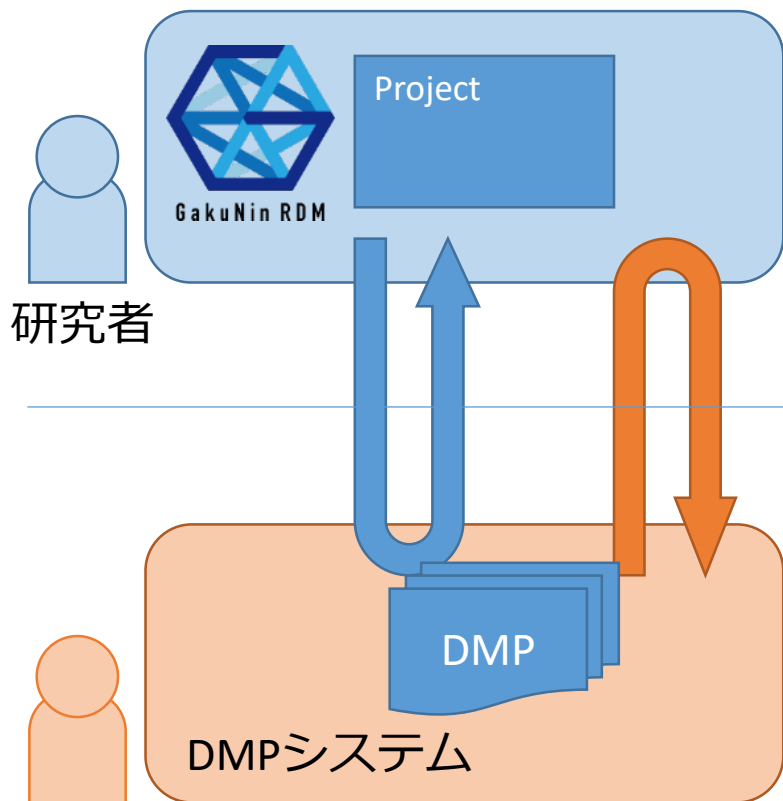
- ・ データのインパクトを評価
- ・ 共同研究の可視化
- ・ データ人材の発掘



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

開発の検討状況： NII RDCにおけるDMPシステムの位置づけ

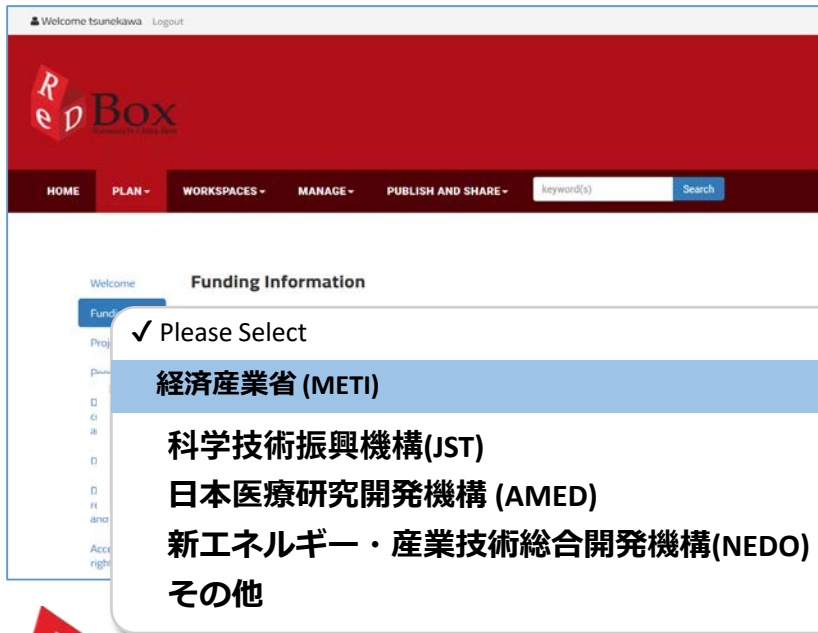


研究データ管理職

- Project Investigator
- Research Office
- Data Librarian

- 研究者が利用する場合
 - GakuNin RDMのサブシステムとして利用し「自分の研究に合った最適なプロジェクトを構成するツール」として使用
- 研究データ管理職が利用する場合
 - DMPを組織的に集約・管理するDBとしてメインで利用
 - DMPと紐づいたGakuNin RDMプロジェクトから研究データ情報を収集することで、データ資産を記録

開発の検討状況：「ReDBoX2.0」をベースとした開発の検討



The Queensland Cyber Infrastructure Foundation (QCIF)とシドニー工科大学が開発をリードするOSSのDMP管理ソフトウェア

NII RDCの要素システムとして拡張を検討

1. 学認への対応
2. GakuNin RDMとの連携
 - 作成したDMPからGakuNin RDMのプロジェクトを自動作成
3. 再現性基盤との連携
4. 国内研究費助成機関への対応
 - JST、NEDO、AMED等が求めるDMPの対応
5. ダッシュボード機能の開発



<https://www.redboxresearchdata.com.au/>