

研究データ管理から研究データ利活用へ ～広島大学と中国四国コンソーシアムによる 研究データエコシステムの構築～

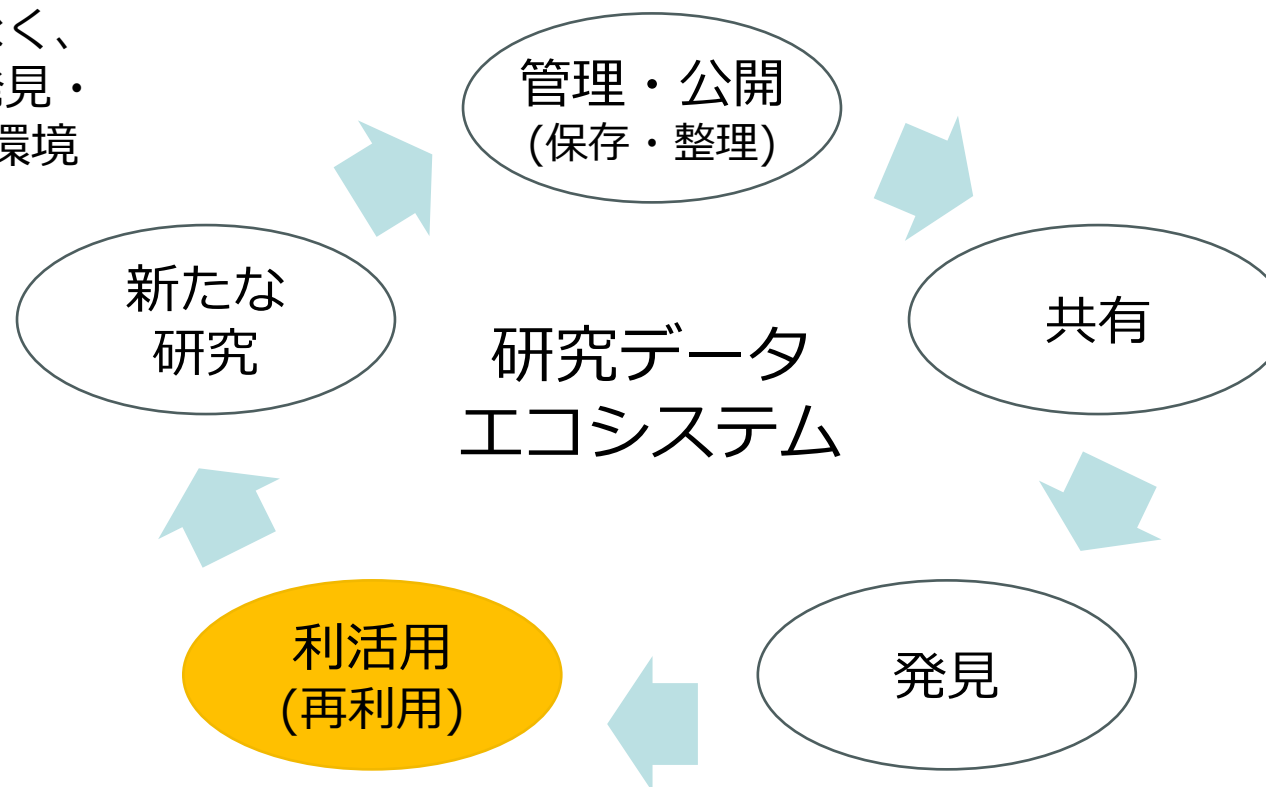
西村 浩二

広島大学 副学長 (情報担当)
情報メディア教育研究センター

神戸大学データ管理・利活用シンポジウム

研究データ管理から研究データ利活用へ

管理・公開は目的ではなく、
研究者が安全に共有・発見・
利活用 (再利用) できる環境
をつくるための手段。

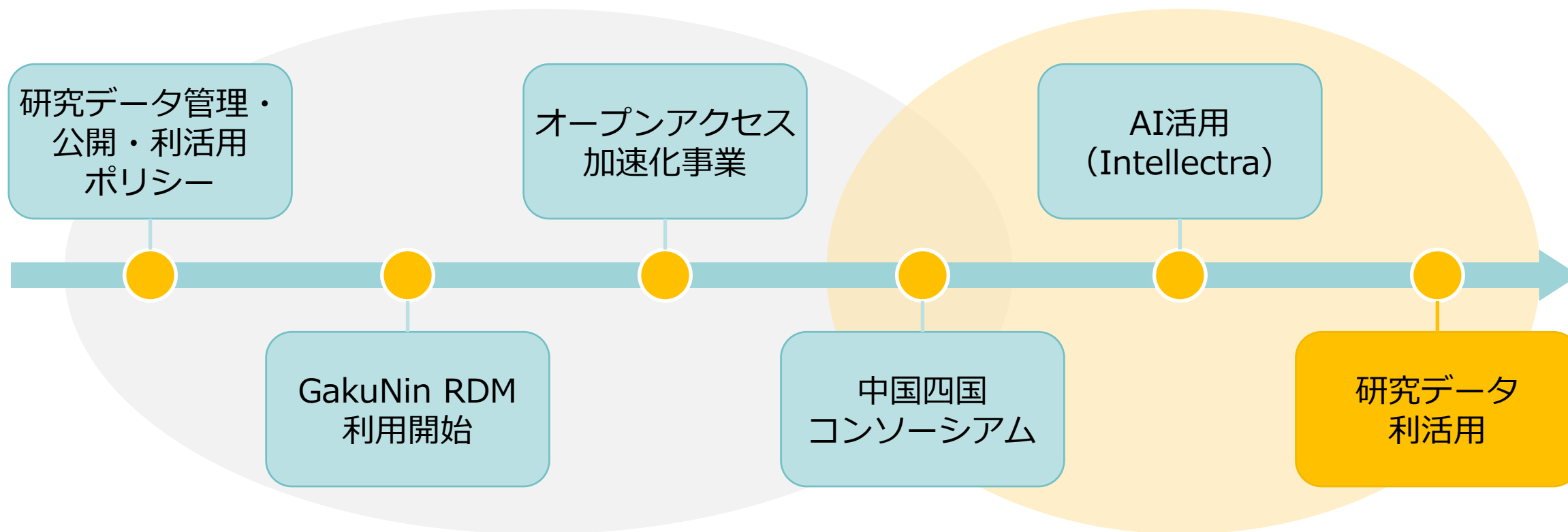


基盤整備

研究者支援

地域連携

広島大学における研究データ管理の歩み



基盤整備から研究者支援・地域連携へ、そして利活用へと段階的に発展

OA加速化事業による研究データ基盤整備

データレイク

研究データをはじめとする学内の多様なデータを集約し、後から探索・再利用できる状態で蓄積する。
・実効容量：1PB超

GakuNin RDM

研究データを安全に保存・共有する共通の入口。
学認により学外の共同研究者とも共有できる。

Nextcloud

研究者が普段どおりファイルを扱える拡張ストレージ。
大容量データの実務的な受け皿。

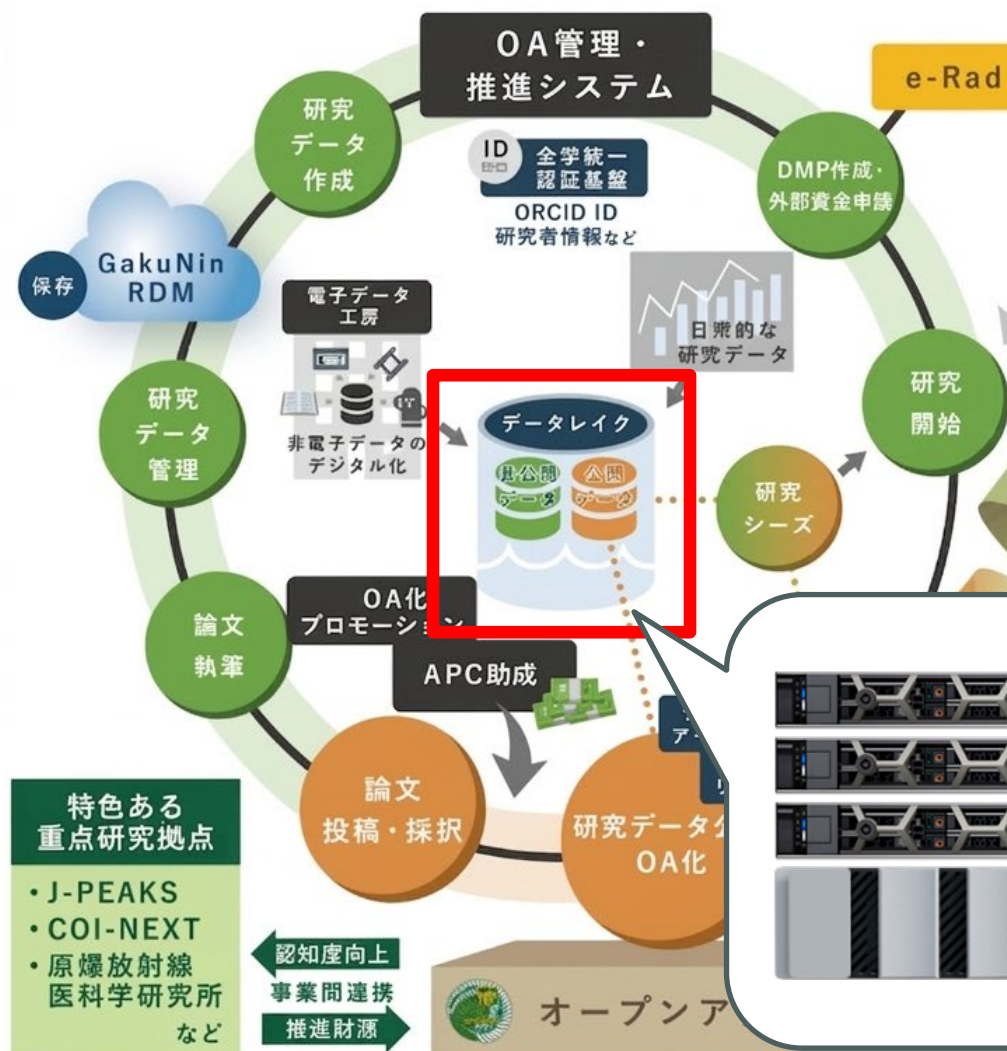
拡張ストレージ

標準容量では足りない研究データを継続的に保管し、公開・利活用につなげる。
・提供容量：1TB (教員)、10GB (職員)

基盤は“保存するため”ではなく、共有・発見・再利用を可能にするために整備する

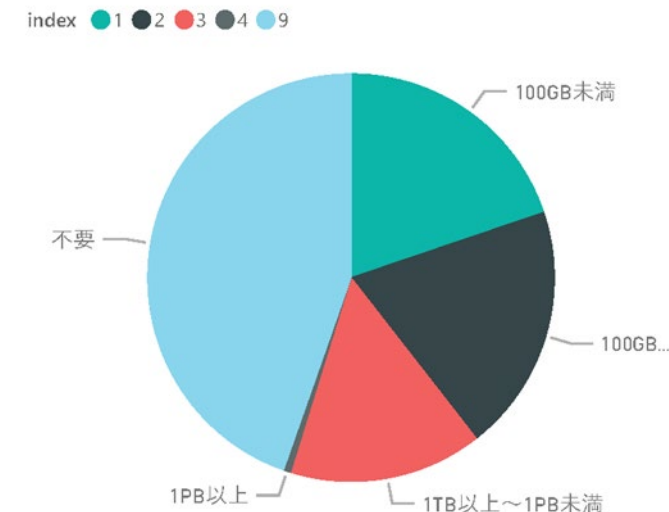
研究者が意識せずに使える状態にすることが基盤整備のゴール

データレイクの容量設計



Q6.1. 保管したい研究データの規模は？ (2019年度アンケートより)

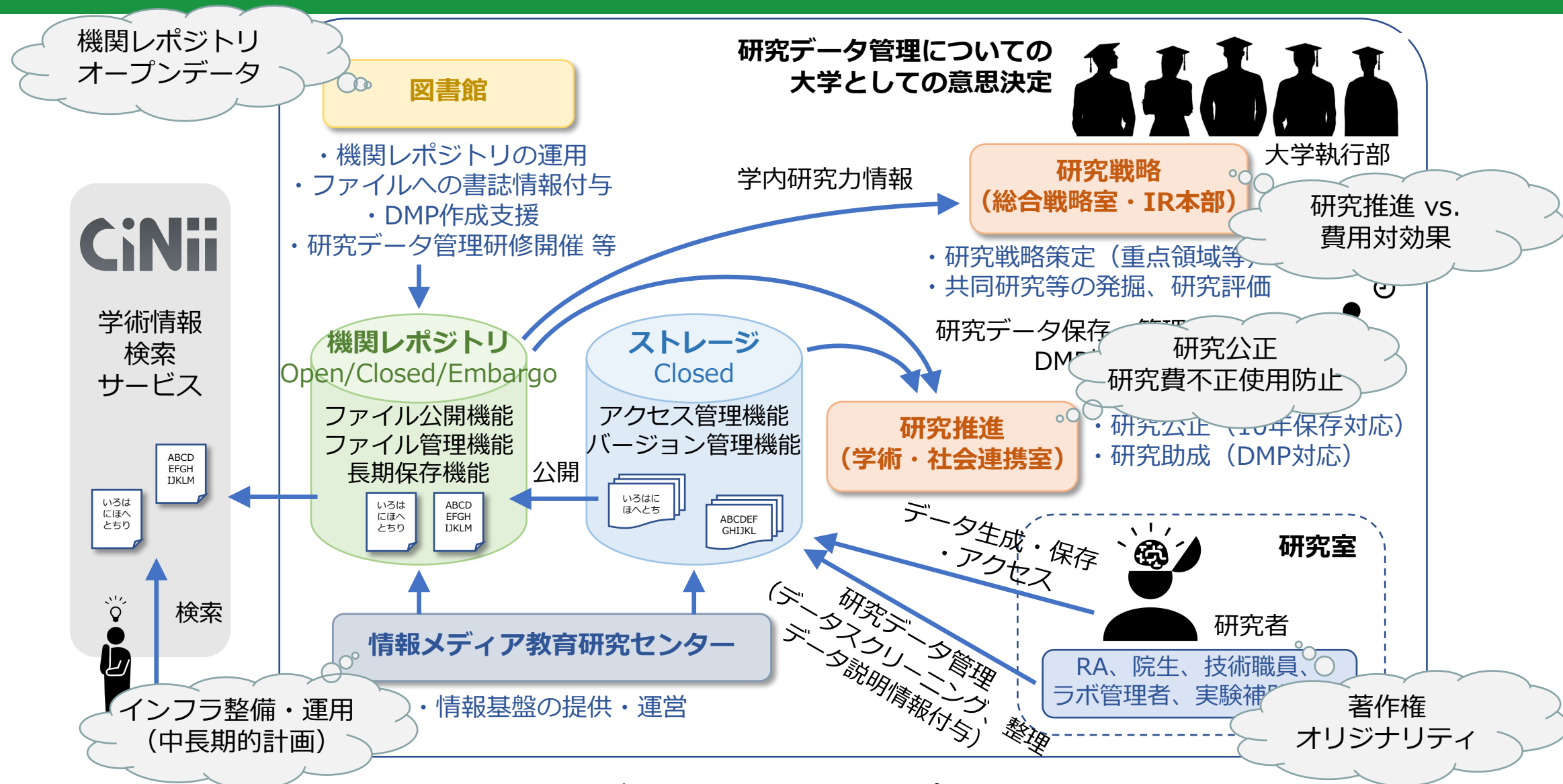
1. 100GB未満 [100GB]
 2. 100GB以上～1TB未満 [500GB]
 3. 1TB以上～1PB未満 [5TB※]
 4. 1PB以上 [5TB※]
 9. 不要
- [] を各カテゴリの代表値とすると
GRDMに約1.1PBが必要
※作業中データは別の場所に保存されると想定



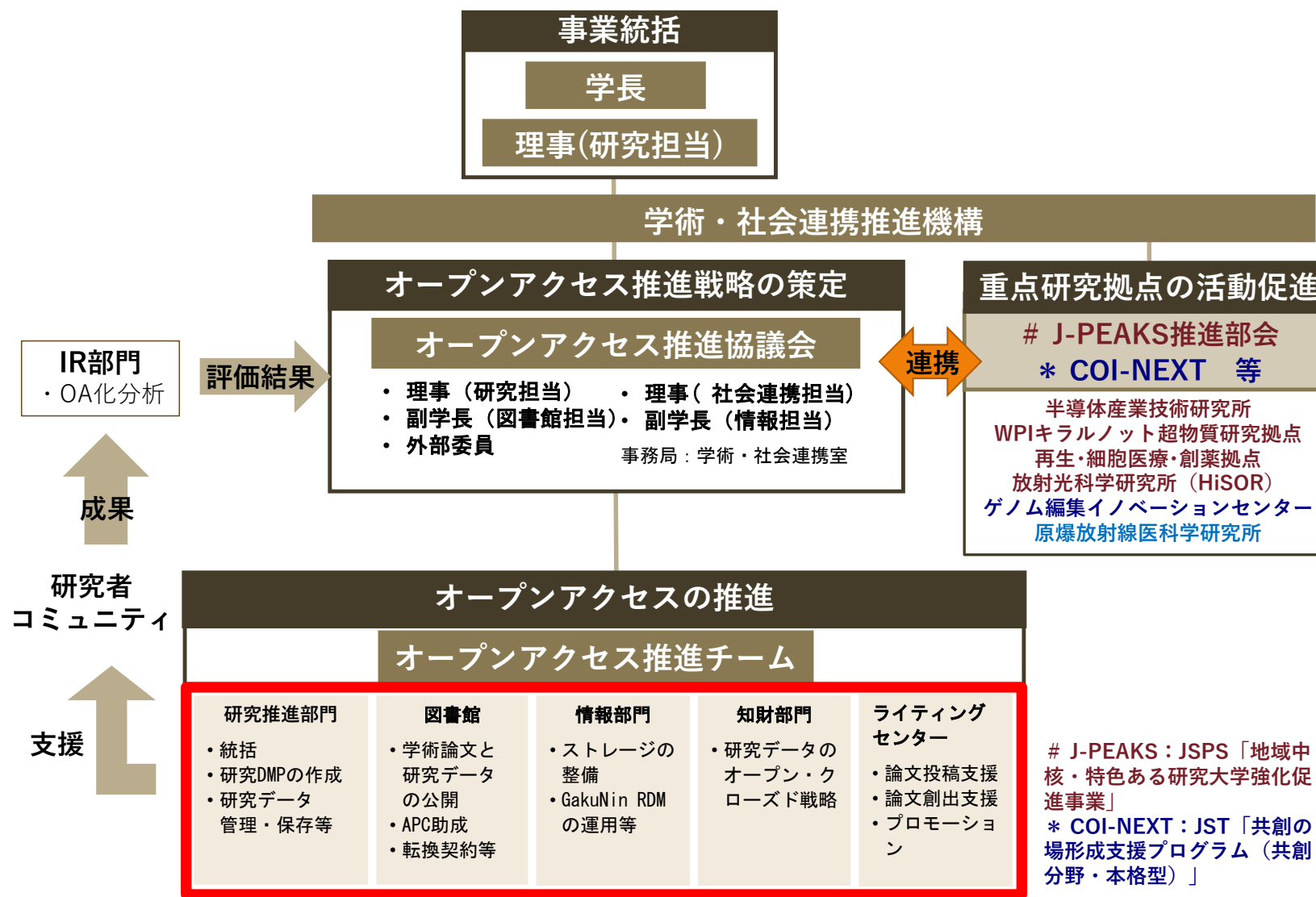
- 運用管理サーバ
 - PowerEdge R660 ×3
 - CPU: インテル® Xeon® Gold 6426Y 2.5G
 - Memory: 32GB×8
 - HDD: 600GB×4
- 大容量高速データストレージ
 - NetApp AFF C30 ×1
 - 物理容量: NVMe 30.7TB×24 (736.8TB)

- ✓ Nextcloud経由でGRDMの拡張ストレージとして利用
- ✓ 重複排除/圧縮処理により、2倍程度の実効容量を見込む

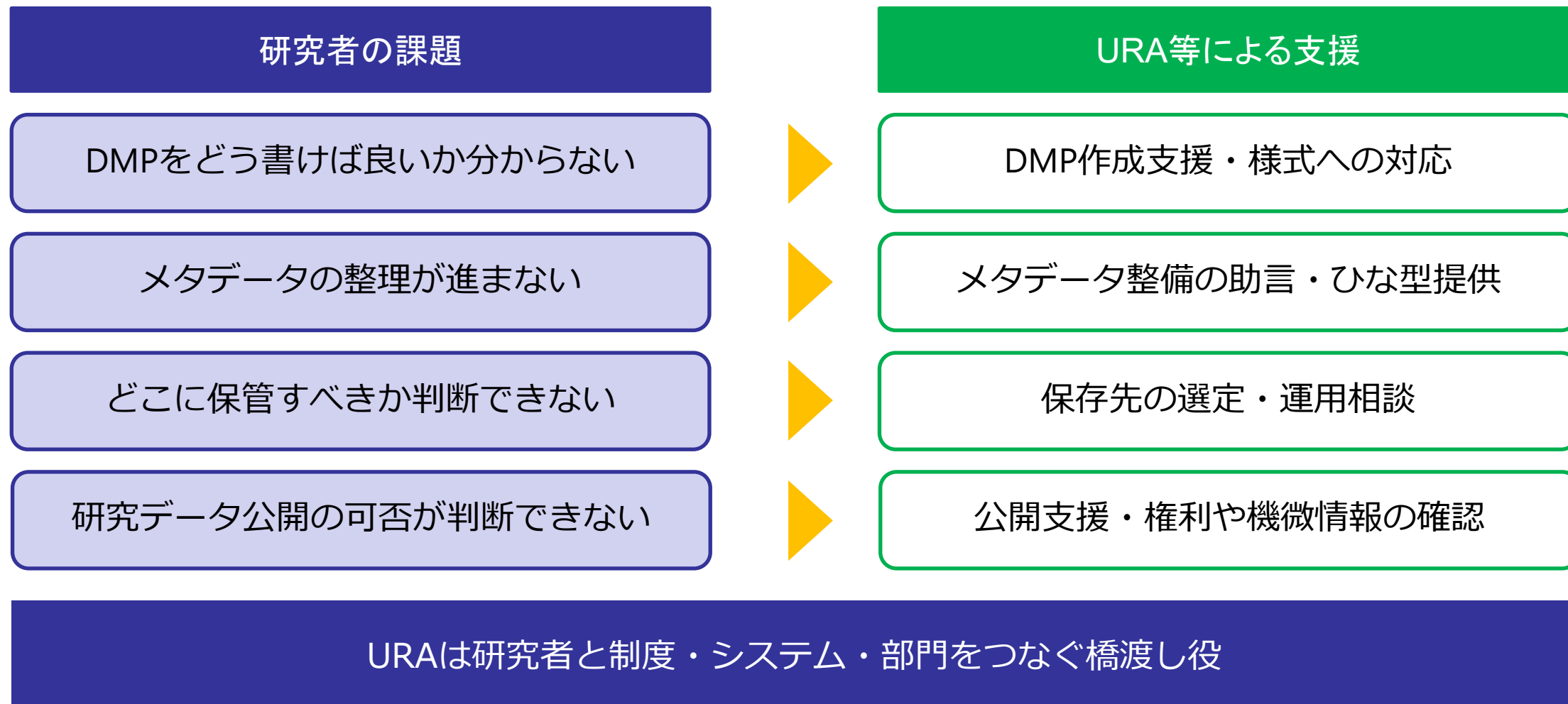
基盤だけでは研究データ管理は定着しない



OA加速化事業の組織体制



URAを核とした研究者支援



中国四国コンソーシアム – 地域で支える研究データ基盤



広島大学

単独の大学では担いきれない研究データ管理を、地域の大学が共同で支える仕組み。

ポリシー共有

各機関の状況に応じた研究データ管理ポリシーの策定を相互に支援し、地域全体の水準をそろえる。

GakuNin RDM活用

構築・運用のノウハウを共有し、機関を越えた研究データの共有を実現する。

情報交換コミュニティ

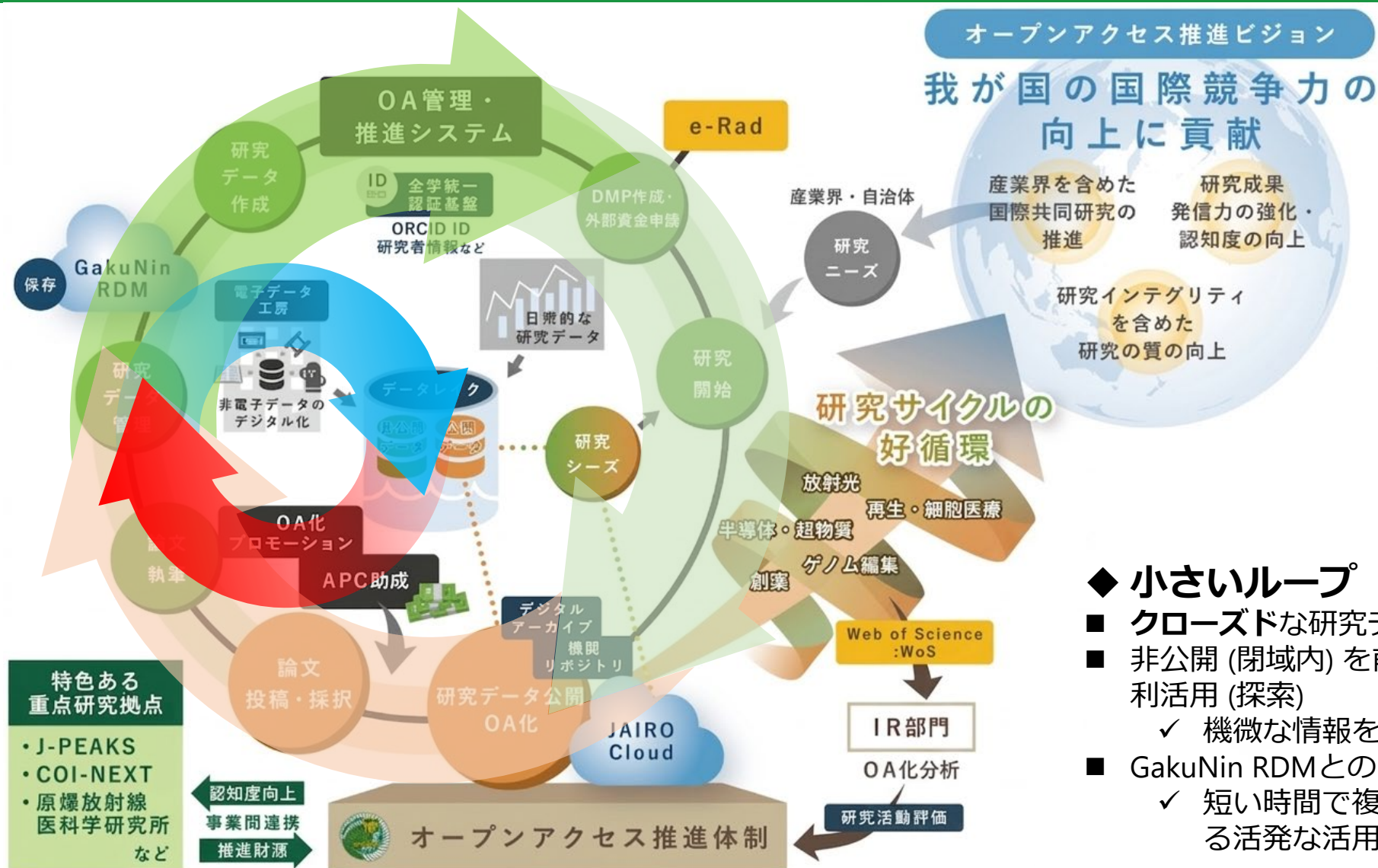
勉強会・シンポジウムを通じて、担当者(URA等)が孤立せず学び合える場をつくる。

AI活用と研究支援

文献調査システムの共同利用など、研究支援そのものを高度化する取り組みへと広げる。

地域連携を“一緒に整備する”から“一緒に活かす”段階へ

管理した研究データをどう生かすか



◆ 大きいループ

- オープンな研究データエコシステム
- 公開を前提とした環境での研究データの利活用
 - ✓ 根拠データのオープンアクセス化
- CiNii Research → GakuNin RDM → JAIRO Cloudの間でグルグル
 - ✓ 1サイクルの時間が長い
 - ✓ 原則として1サイクルで完結 (次のサイクルに繋げることが課題)

◆ 小さいループ

- クローズドな研究データエコシステム
- 非公開 (閉域内) を前提とした環境での研究データの利活用 (探索)
 - ✓ 機微な情報を含む可能性
- GakuNin RDMとの間でグルグル
 - ✓ 短い時間で複数のサイクル (共同研究等における活発な活用に期待)

GakuNin RDMとIntellectraによる研究データ利活用



研究者がアップロードした研究データが、そのまま探索・再利用の対象となる。

閉環境での探索

AIの活用

非定型ドキュメントの活用

共同研究の支援

新たな知識発見

管理された研究データが、研究者にとって“探せる・使える”資産になる

まとめ

1

基盤を整える

GakuNin RDM・拡張ストレージ・データレイクにより、研究データを安全に扱える環境を用意する。

2

人と組織をつなぐ

URA・図書館・情報部門・研究推進部門が連携し、研究者の実務を支える。

3

管理から利活用へと進む

地域連携とAI活用により、研究データエコシステムを実現する。

研究データ管理を、研究者への価値創出につなげる

管理は目的ではなく、利活用を実現するための基盤である