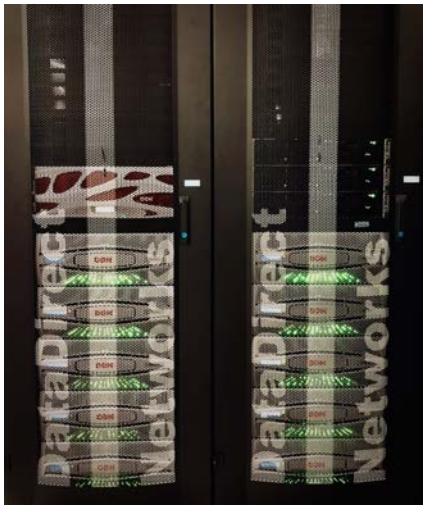
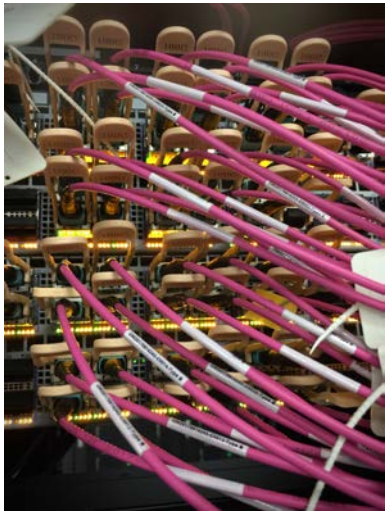
# Gfarm Workshop 2019



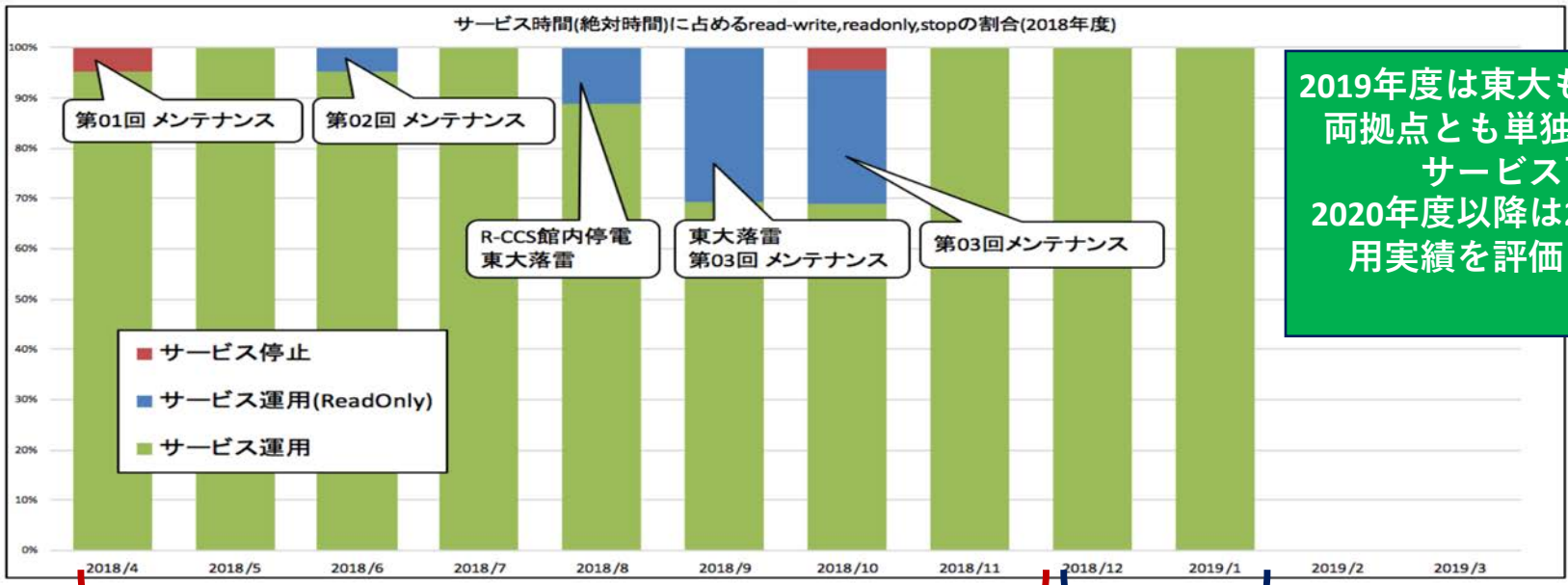


理化学研究所　計算科学研究センター

原田 浩

2019/1/31

# HPCI共用ストレージ現在地：可用性




2019年度は東大も二重化して、両拠点とも単独でReadWriteサービス可能に
2020年度以降は2019年度の運用実績を評価して。。。

片拠点運用時はReadOnlyサービス

R-CCS単独運用時もReadWriteサービス

- 2018年度は、時間割合で99.1%(2019/1/28時点)サービスを提供
- 99.1%のうち、ReadOnlyサービス：6.84%、ReadWriteサービス： 92.3%
  - ReadOnlyの原因はR-CCS 台風による館内停電、東大　落雷による瞬停、第3回メンテナンス
- サービス停止は計画した二回だけ。不意のサービス停止は発生していない
- 2018年12月4日にR-CCS内でもデータ二重化し、R-CCS単独運用でReadWriteサービス提供可能に
- 2018年12月14日開始の第４回メンテナンスでは、ReadWriteサービスを継続
- サイト間データ冗長(二重化)運用の効果は絶大

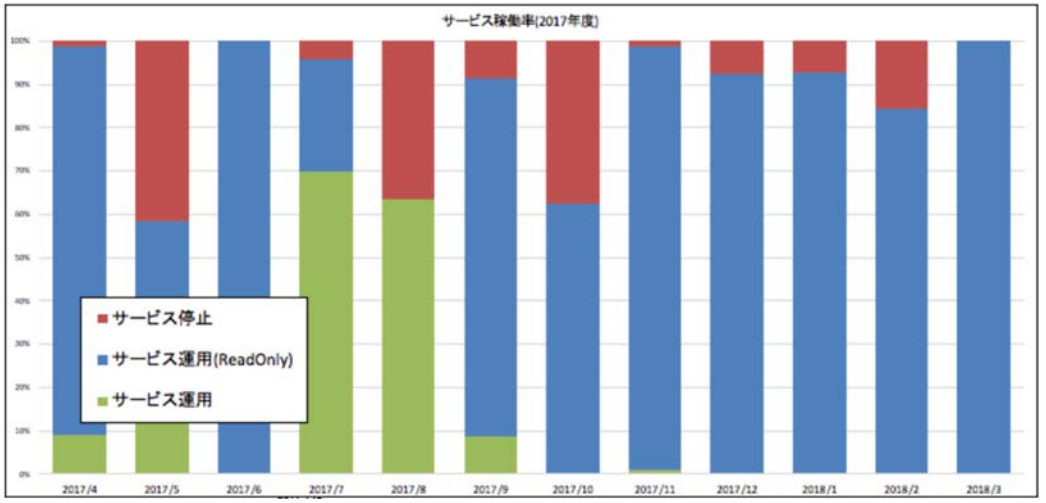
# 決して順調だったわけではない
# 台風にもマケズ、雷にもマケズ
# HPCI共用ストレージはサービス継続

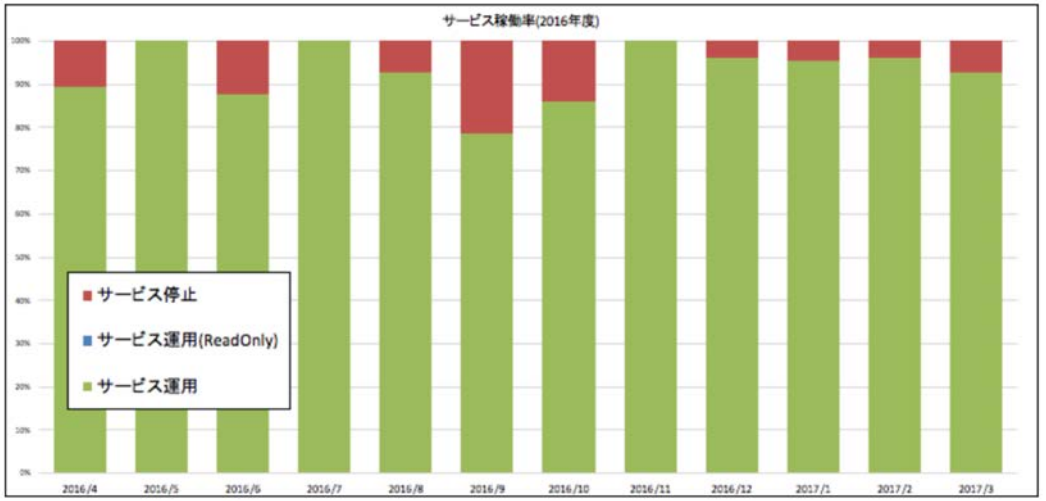
- 8月23日深夜：台風20号の影響でR-CCS設備(電源供給)障害
- 8月31日18時：落雷の影響で東京大学柏キャンパス情報基盤センター停電発生
- 9月26日～10月10日、東大柏、R-CCS、２週続けて週末計画停電のため長期メンテナンス
- サービス停止は、ファイルサーバのドメイン名、IPアドレス変更のための３日間だけ
- データ破損、消失も無し

ネットワーク含むインフラは、落とし穴
ネットワークメンテナンス・障害はサービス停止につながる

# HPCI共用ストレージ過去：可用性

- 2017年度は、機材更新とデータ移行
- 機材更新・データ移行のメンテナンスのためReadOnly運用が多かった
  - 稼働率：86.46%
  - RadOnly稼働率：71.67%
- 2016年度は、旧機材かつデータ1.5倍割り当て時代
  - 稼働率：92.92%
  - RadOnly稼働率：0%
- メンテナンス時はサービス停止
- 特に9月から10月にかけて東大、R-CCSで二週続けて設備点検を実施するので、どうしても長期のサービス停止を伴ってしまった
- 2018年度の稼働率99%超だがReadWrite稼働率は92.3%で2016年度が若干優れる



# HPCI共用ストレージ2019年度：可用性

共用ストレージ

【共用ストレージ】R-CCS設備工事(サービス無停止)のお知らせ（2019年10月ころ)

作成者 理化学研究所 計算科学研究センター、最終変更日2019/01/18

お世話になっております、HPCI共用ストレージ担当です。
理化学研究所　計算科学研究センターでは2019年10月頃に大規模な設備工事を予定しております。
工事期間中は、R-CCS設置の共用ストレージ機器は運用を停止する予定です。
R-CCS設置機器の停止期間中も、東京大学設置機器による片拠点運用を予定しているため、
通常通りHPCI共用ストレージをご利用いただける予定です。
現在、R-CCS設置機器停止に備え東京大学設置のストレージ機器におけるデータ二重化を進めております。
停止期間などの詳細につきましては、決まり次第このページに記載する予定です。
ご協力のほど、よろしくお願い申し上げます。
--
HPCI共用ストレージ担当

- ReadOnly運用時間を削減して、過去最高の稼働率を目指したい
- 2019年度も9月~10月は鬼門
- R-CCS長期設備工事のため2weeks以上機器停止を予定
- 東大片肺運転でサービス継続できそう
  - マスタメタデータサーバの東大運用
  - 東大設置ストレージによるデータ多重化
- 東大の設備点検とR-CCS設備工事が被らないことを祈りつつ！
- 9月10月をしのげれば、さらなる稼働率向上を目指せる
- 2019年度でデータ多重化運用の評価も固まると見込まれる。2020年度以降のデータ多重化方針も決められる

# HPCI共用ストレージ現在地：HDD障害

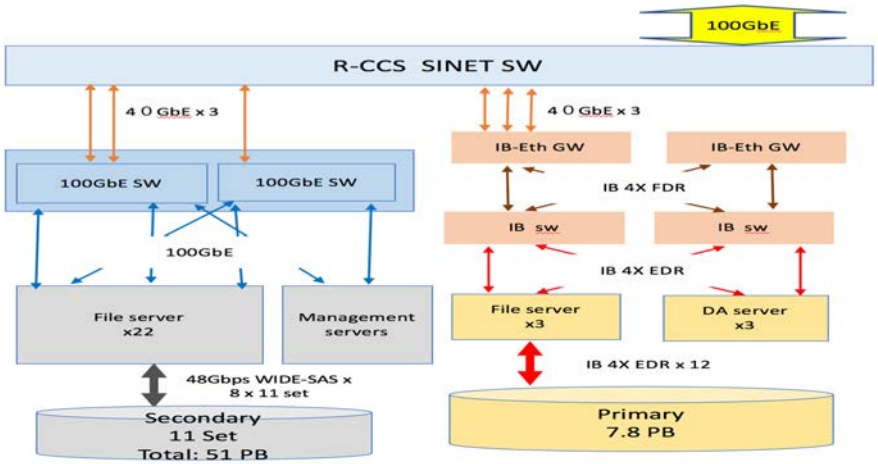
- R-CCSでは一次ストレージ＋二次ストレージの２機種を利用
- R-CCSへのアクセスは、一次ストレージに優先書き込み
- 一次・二次ともHDDのMTBFは500万h超で公称値の２倍以上
- RAID6破損障害は発生していない
- 二次ストレージは定期的に全ディスクをS.M.A.R.Tで状態監視
- MTBF500万hで2P+8DのRAID6が破壊される→MTBF500万hのHDD10本のうち、任意の3本が25H以内に同時に壊れる→7Phに一回くらい？
- このまま続いてね
- 第一世代のHDDより高信頼
- まだ運用時間が短期間だけど。。。

**一次ストレージ：公称MTBF 250万h、12TB**

| 運用期間 | 運用時間(h) | 運用HDD数 | 障害HDD数 | MTBF(10000h) |
|---|---|---|---|---|
| 2018/4～2019/1 | 7344 | 800 | 1 | 587.5 |

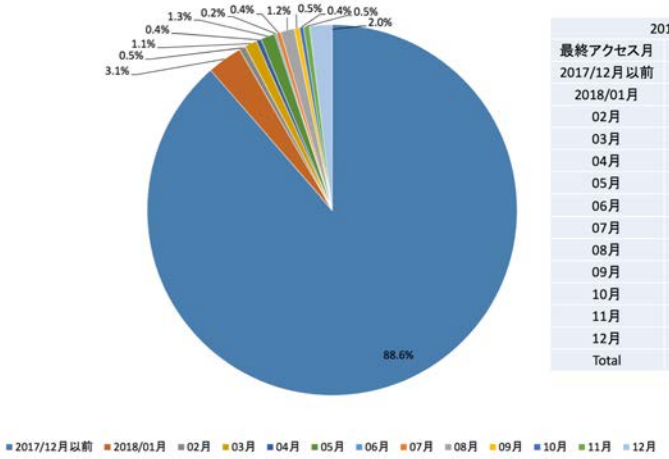
**二次ストレー ：公称MTBF 250万h 、10TB**

| 運用期間 | 運用時間(h) | 運用HDD数 | 障害HDD数 | MTBF(10000h) |
|---|---|---|---|---|
| 2017/10～2017/12 | 2208 | 2320 | 12 | 547.9 |
| 2018/1~2019/1 | 9504 | 6380 | | |



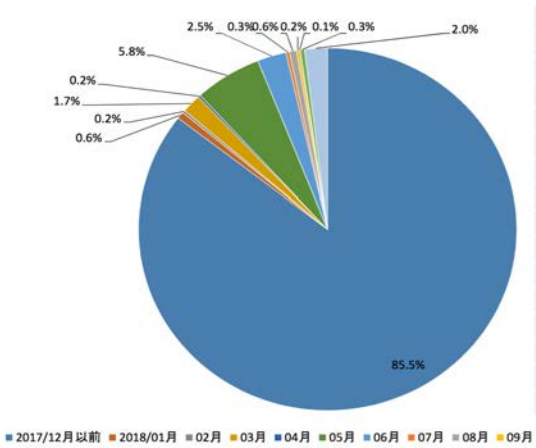
# HPCI共用ストレージ現在地：ユーザの用途

i. 最終アクセス月毎のファイル分布（総容量 TB：複製無）




| 2019/1/7時点 | | |
|---|---|---|
| 最終アクセス月 | 総数 | 比率 |
| 2017/12月以前 | 6,704 | 88.6% |
| 2018/01月 | 232 | 3.1% |
| 02月 | 41 | 0.5% |
| 03月 | 81 | 1.1% |
| 04月 | 33 | 0.4% |
| 05月 | 95 | 1.3% |
| 06月 | 12 | 0.2% |
| 07月 | 29 | 0.4% |
| 08月 | 91 | 1.2% |
| 09月 | 35 | 0.5% |
| 10月 | 27 | 0.4% |
| 11月 | 39 | 0.5% |
| 12月 | 148 | 2.0% |
| Total | 7,568 | 100.0% |

h. 最終アクセス月毎のファイル分布（総数）




| 2019/1/7時点 | | |
|---|---|---|
| 最終アクセス月 | 総数 | 比率 |
| 2017/12月以前 | 42,021,080 | 85.5% |
| 2018/01月 | 311,536 | 0.6% |
| 02月 | 110,303 | 0.2% |
| 03月 | 813,218 | 1.7% |
| 04月 | 101,133 | 0.2% |
| 05月 | 2,839,110 | 5.8% |
| 06月 | 1,218,223 | 2.5% |
| 07月 | 146,902 | 0.3% |
| 08月 | 289,914 | 0.6% |
| 09月 | 118,618 | 0.2% |
| 10月 | 51,933 | 0.1% |
| 11月 | 167,946 | 0.3% |
| 12月 | 960,824 | 2.0% |
| Total | 49,150,740 | 100.0% |

- 明らかに研究成果などのアーカイブ・バックアップ用途が主
  - 1年以上寝ているデータが、容量、数とも85%超
  - 2017年度の長期ReadOnly運用の影響？
  - 今後もアーカイブ・バックアップ主体の用途は変わらないのでは
- 当初HPCI共用ストレージが整備された時の構想とは異なっている
  - 当初の構想では。。。。
- アーカイブ・バックアップが主な用途なら、次のターゲットは研究成果・データの二次利用促進？

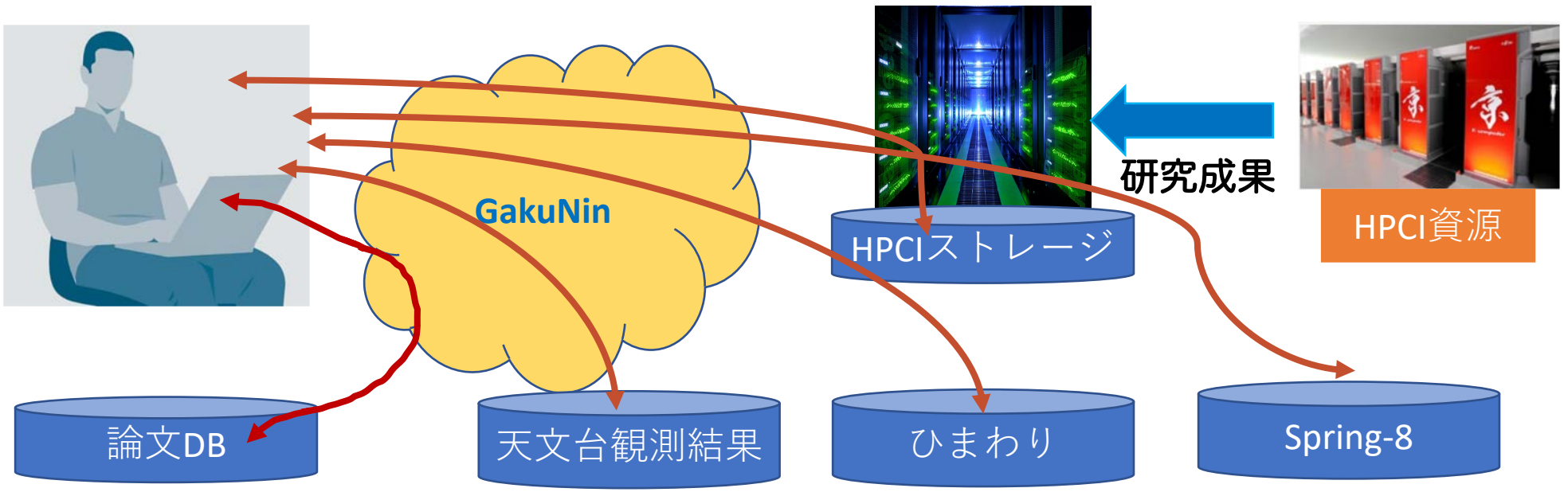
# HPCI共用ストレージ現在地：ユーザの用途

- 研究室単独ではPB級のデータ管理・保存・公開に対応できない(マンパワー)
- PBのストレージをパブリッククラウドに保存できる資金がない(資金)
- パブリッククラウドはセキュリティ問題により、データを保存できない(セキュリティ)
- クラウドに成果を保存するとクラウドベンダーにロックインされるだけ(継続性)
- 研究者が片手間に管理できるサイズを超えている
- 外部資金で成立している研究プロジェクトの成果は、外部資金が切れたら管理できない
- 10年もデータを保管できない
- 研究者が移動・退職してもデータを管理できる仕組み
- シミュレーションデータだけではなくて、ソフトウェアやデジタル画像などもコンテンツとして管理・活用できる仕組みが必要
- データサイエンスの立場で言えば、最終的に“コンテンツ“を所有し、二次利用可能な体制を持っているところが勝つ
- データ管理・データ活用のための戦略的かつ大規模な取り組みが必要

# HPCI共用ストレージ：二次利用促進に向けて

- NII オープンサイエンス基盤研究センター ： Research Data Cloud で研究成果を公開・二次利用促進のためのサービスを提供開始
- GakuNin weko3
  - https://www.nii.ac.jp/event/upload/libfair2018_forum2_3.pdf
  - 成果データ活用のためのポータルを提供
  - NIIのクラウドから無料で利用できる(容量などはミニマムサービス)
  - メタデータと識別子を登録可能
  - 研究成果活用向けリポジトリとして、 ローカルストレージ・AWS W3などを指定可能
  - **公開データの保存場所として、HPCI共用ストレージを使えないか**
- GakuNin RDM
  - コミュニティ内での成果・データの管理・共有
  - NIIのクラウドから無料で利用できる(容量などはミニマムサービス)
  - 検索キーワードなども登録可能(スキーマはカスタマイズ可能)
  - データの保存場所は、基本的にローカルストレージ・不特定多数からのアクセスは考えていない
- CiNii -- 論文検索サービスでお馴染み

# HPCI共用ストレージ：二次利用促進に向けて




RCOS

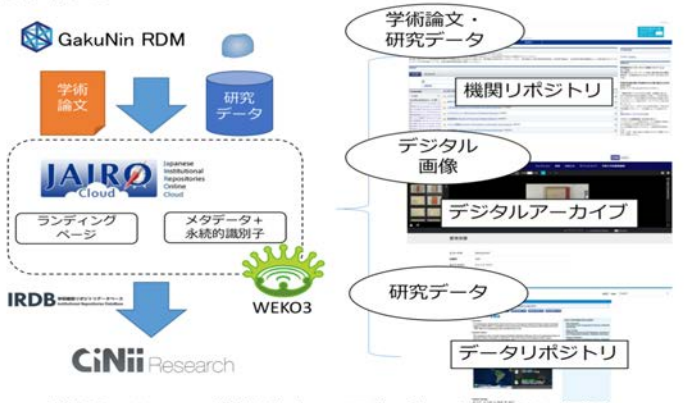
## WEKO3はどうなるのか？

JC（WEKO2)の機能を基礎にWEKO3を開発

- **簡便セルフアーカイブ機能**
  - GakuNin RDMとの連動機能の提供
- **カスタマイズ可能な公開ワークフロー機能**
  - WEKO2のワークフロー機能を強化
  - カスタマイズ機能の追加
  - 登録・公開ワークロード軽減機能の提供
- **コレクション機能**
  - 特定のコレクション毎にアイテムを表示・提供する機能
  - WEKO2のメタデータ項目定義・配信機能を強化し、分野に応じたメタデータの出力
- **外部システム連携機能（WEB API）**
  - WEKO3からの情報取得が可能となるAPIの提供

RCOS

## 次期JC




機関における学術論文・研究データ公開を支援

ありがとうございました。