



Title	大阪大学オープンサイエンスシンポジウム：オープンサイエンス時代の研究基盤と人材育成：日本における実装と展望
Author(s)	Ganguly, Raman; Gergely, Éva; 富浦, 洋一 et al.
Citation	
Version Type	VoR
URL	https://doi.org/10.18910/101962
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

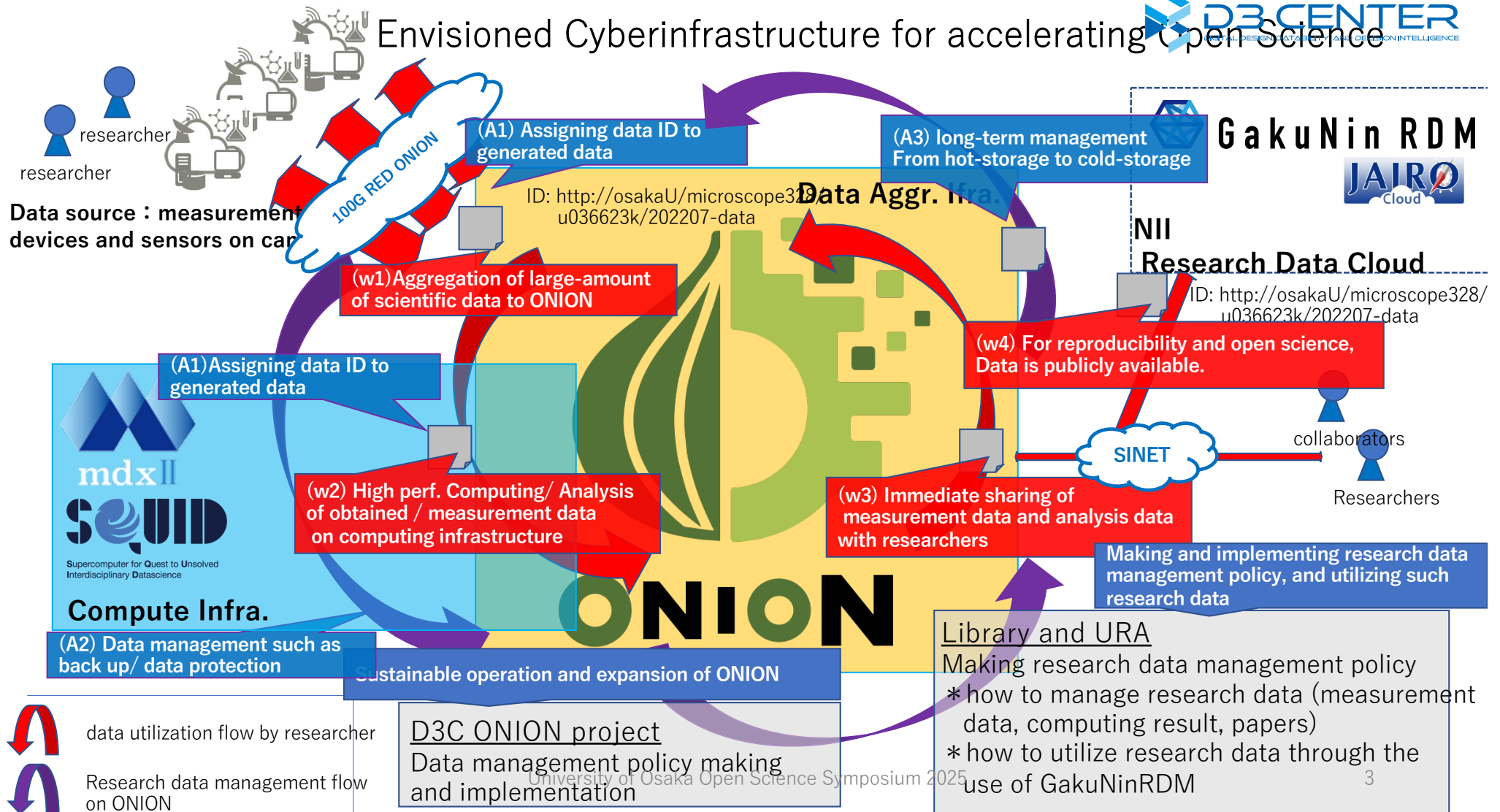
<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

オープンサイエンスを加速する 大阪大学の計算基盤・データ基盤の現状と課題

大阪大学D3センター
先進高性能計算基盤システム研究部門 /
高性能計算・データ分析融合基盤協働研究所
伊達 進

Envisioned Cyberinfrastructure for accelerating Open Science



計算基盤とデータ基盤の連携・融合が研究管理に必須なのか

- 学術研究で扱うデータはほとんどが”デジタル(digital)”データとして取得され、計算機で解析・処理され、保存・管理・公開される。
- とりわけ今日では、学術研究の発展には、**AI(Artificial Intelligence)技術の適用・応用・開発が必要不可欠**となりつつある。
 - **デジタルツイン (Digital Twin)**：実世界のデータから**高精度に現実空間を再現(シミュレート)**すると同時に、AI技術により高精度に実世界の予測・予見・最適化を行う



- 高精度な実世界のシミュレーション、高精度なAIの実現には、大容量データを収容するデータ基盤および、その大容量データを解析・分析処理する計算基盤が不可欠
- このデータ基盤と計算基盤の上で創出される研究データを、学術研究基盤として責任ある形で管理し、研究の再現性(Reproducibility)を保証する仕組みづくりが急務

SQUID since May 2021



Supercomputer for Quest to Unsolved
Interdisciplinary Datascience

- クラウド連動型高性能計算・高性能データ分析用スーパーコンピュータ
(Supercomputer for Quest to Unsolved Interdisciplinary Datascience)
 - 総理論演算性能 16.591 PFlops



最新 3 種混合プロセッサ・アクセラレータ搭載！

SQUIDは、2021 年 4 月に発表されたばかりの Ice Lake 世代のプロセッサをはじめ、最新 GPU アクセラレータ NVIDIA HGX A100 8 GPU ボード、最新ベクトルプロセッサ NEC SX-Aurora TSUBASA Type 20A を搭載しています。汎用 CPU ノード群は、Ice Lake 世代のプロセッサを搭載する国内最大級となります。

SQUID システム構成



汎用 CPU ノード群

1,520 ノード x 理論演算性能 5,837 TFLOPS **8.871 PFLOPS**

プロセッサ Intel Xeon Platinum 8368 (Ice Lake / 2.40 GHz 38コア) 2基

主記憶容量 256 GB

GPU ノード群

42 ノード x 理論演算性能 161.836 TFLOPS **6.797 PFLOPS**

プロセッサ Intel Xeon Platinum 8368 (Ice Lake / 2.40 GHz 38コア) 2基

主記憶容量 512 GB

GPU NVIDIA HGX A100 8 GPU ボード (Delta)

ベクトルノード群

36 ノード x 理論演算性能 25.611 TFLOPS **0.922 PFLOPS**

プロセッサ AMD EPYC 7402P (2.8 GHz 24コア) 1基

主記憶容量 128 GB

Vector Engine NEC SX-Aurora TSUBASA Type 20A 8基

ノード間接続

ノード間接続 Mellanox InfiniBand HDR (200 Gbps)

データ集約基盤 ONION

S3 対応並列ファイルシステム 21.2 PB

ファイルシステム DDN EXAScaler (Lustre)

HDD 20.0 PB

SSD 1.2 PB

S3 対応オブジェクトストレージ 500 TB

オブジェクトストレージ CLOUDIAN HyperStore

HDD 500 TB

SQUID+ONION 5つのチャレンジ



データ集約基盤ONION

- Osaka university Next-generation Infrastructure for Open research and open Innovation
- 世界最高水準の基礎的、基盤的研究や学際融合研究が生み出す多様な知の創出と深化に寄与すべく、総合大学である大阪大学内で創出された「利用可能な超大量データを将来に渡る持続可能性を保持しつつ責任をもって活用」可能にするとともに、新たな社会的価値の創出を目指した「産学共創」「国際共同研究」のための学内外でのデータ利活用を支援するデータ集約基盤
- スーパーコンピュータSQUIDの調達に合わせてPoC (Proof-of Concept)実装として試験的に導入

- 本格的に大規模なONIONを整備してしまうと、調達主目的のスーパーコンピュータの規模を大幅に縮減してしまう。
- 構想したONIONが有用でなく、利用されない場合もありえる

特徴2. クラウド連動・連携機能

高負荷状態時にオフロード
新製品の早期実戦配備

民間クラウド(IaaS)事業者

ODINS網を活用した
学内部局からのデータ集約

クラウドストレージを活用した
海外研究拠点とのデータ共有

特徴4. データ集約環境

民間クラウド(ストレージ)事業者

データ集約基盤 ONION

大容量高速ストレージ
20+1 PB (HDD+SSD)

サイバーインフラ高度化への課題(1)

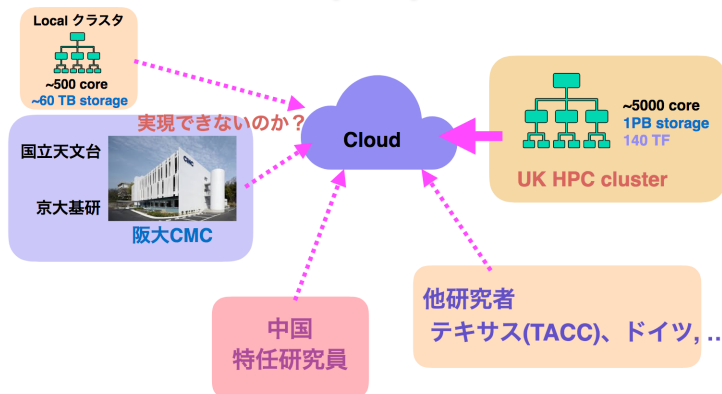
- ・ **学術研究の広域化・グローバル化、産学共創への期待**
 - ・ 共通の課題解決にむけて世界の研究者と協働する国際共同研究



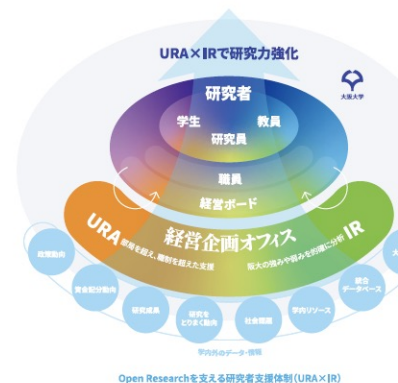
地理的に分散する計算資源、データ資源を駆使して、
研究開発できる環境

Open Research

阪大の国際共同研究促進プログラム



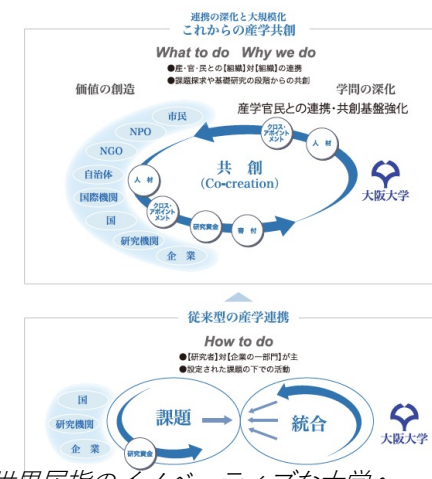
研究者のときめきと自由な発想による
学術研究を基軸とし、専門分野を超え、
広く世界と協働する新たな知の創出を目指す
「Open Research」を推進します。



Open Researchを支える研究者支援体制(URA x IR)

Open Innovation

「産学連携から産学共創へ」をコンセプトに掲げ、
社会のニーズに基づく基礎研究の課題を発掘し、
新たな社会的価値の創出につながる
「Open Innovation」に挑戦します。



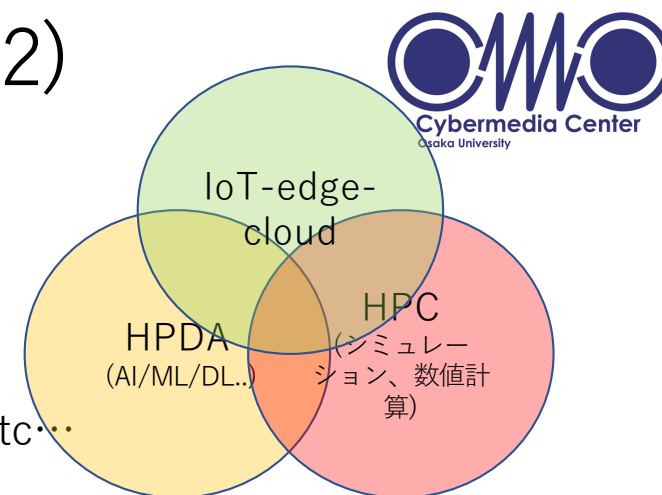
長峯健太郎, “理学研究とクラウド利用のニーズ:理論宇宙物理学の例”,
Cyber HPC Symposium 2019, Mar. 2019. 発表スライドより

University of Osaka Open Science Open System 2021 社会変革に貢献する世界屈指のイノベティブな大学へ、
<https://www.osaka-u.ac.jp/ja/oumode/OUvision2021/y66s8j>

サイバーインフラ高度化への課題(2)

• HPDA分野の計算ニーズの急速な高まり

- AI/MLを活用するデータ駆動型サイエンス
- 各種IoTセンサを利活用したビッグデータ解析
- HPC for AI, AI for HPC
 - データ同化シミュレーション, AIによる計算結果の判定, etc...



多様なデータをスーパーコンピューティング環境に“集約”できるデータ基盤
= 学術研究のデータフローを妨げないデータ基盤
= 利用者のデスクトップ環境、クラウド、スパコン環境を問わず
データ（計測データ、計算結果、等）を簡易に移動・管理できる環境

サイバーインフラ高度化への課題(3)



• 研究データ保存, Reproducibilityに向けた動きの活発化

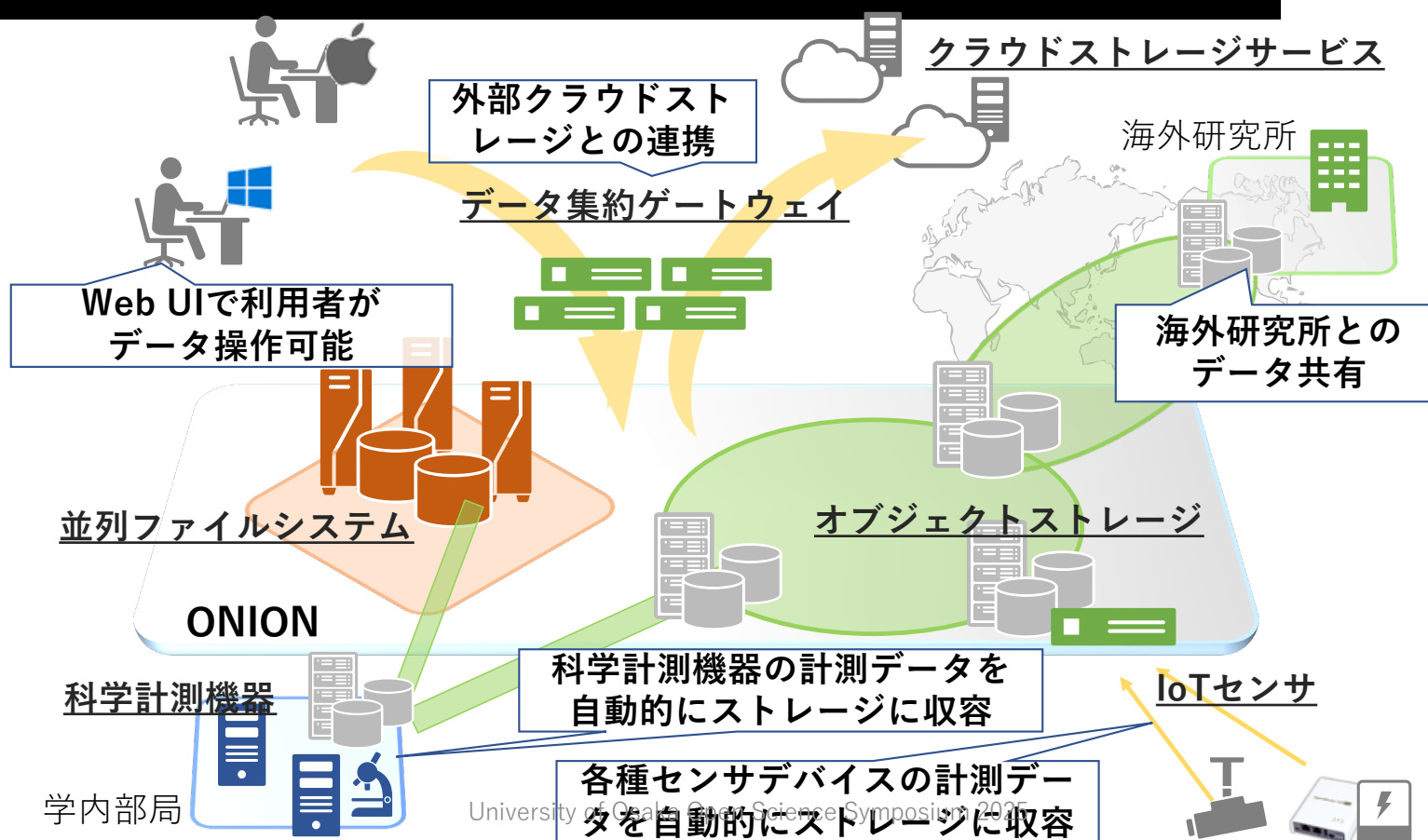
- 公正な研究活動の推進にむけて、研究活動に伴い作成・取得した研究データの保存期間および管理方法等についての基準を定めたガイドライン策定
- 様々な科学データ計測機器がキャンパス内に存在しており、そのデータ管理・移動が課題
- 大容量・大規模データを取り扱うSCセンターとしての役割・責任



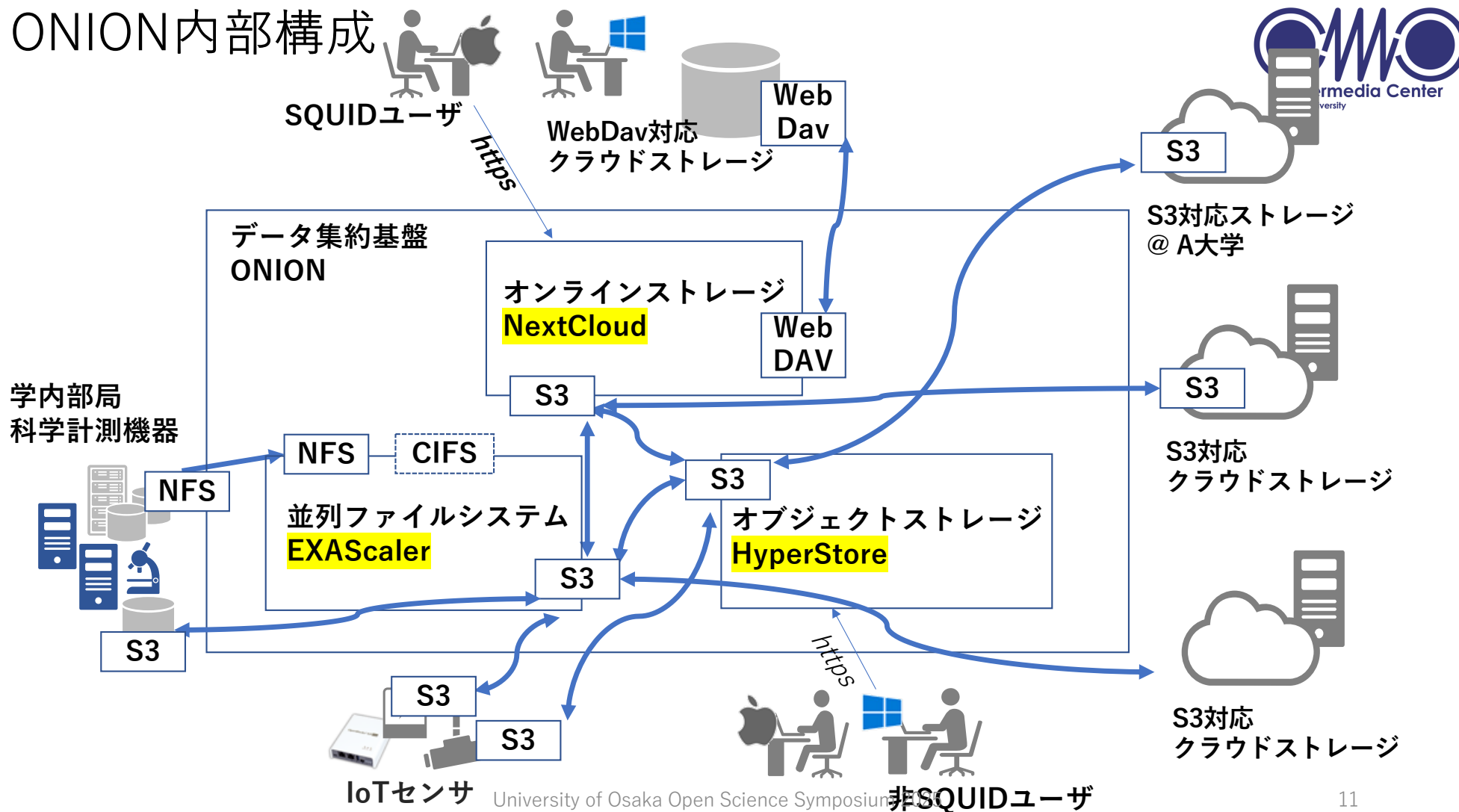
近い将来、SCセンターとしてデジタル化された研究データを適正に管理するためのデータ基盤を提供する必要がある？

ONIONの概要

3種のストレージソリューションを相補完的に連動



ONION内部構成



ONION-OUKA

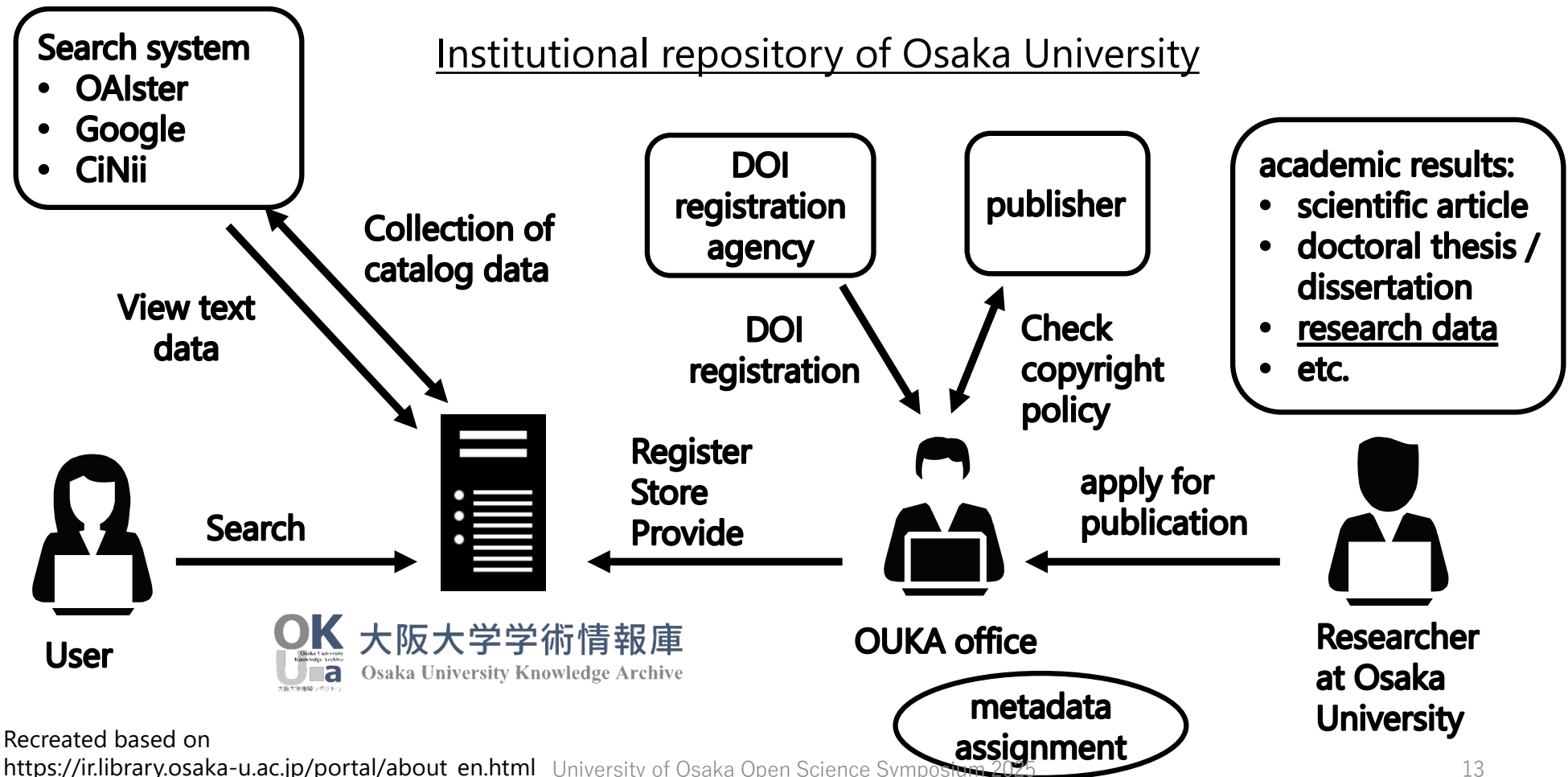
データ公開プロセスを促進する仕組み

田主 英之, 山下 晃弘, 細見 岳生, 並木 悠太, 甲斐 尚人, 松浦 かな, 伊達 進, “研究データ管理を支える学内情報基盤連携の実現に向けて”, 学術情報処理研究, 2023. [[DOI:10.24669/jacn.27.1_98](https://doi.org/10.24669/jacn.27.1_98)]



Courtesy of 田主英之(大阪大学 D3センター高性能計算・データ分析融合基盤協働研究所)

機関レポジトリ OUKA

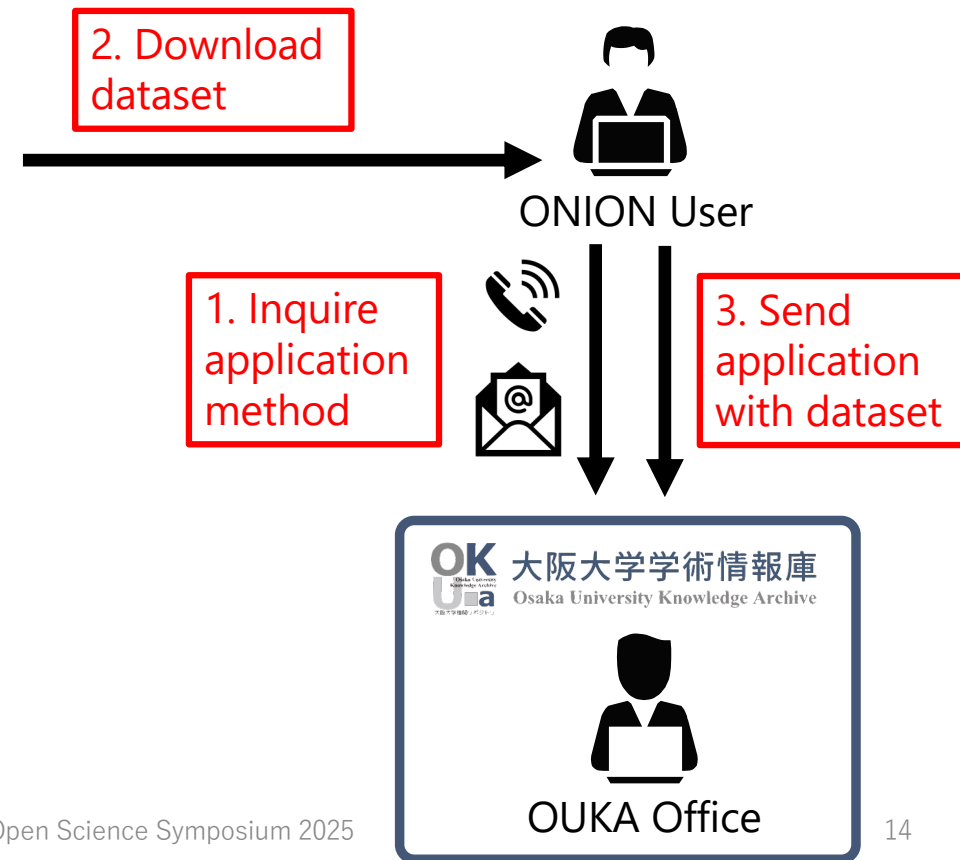
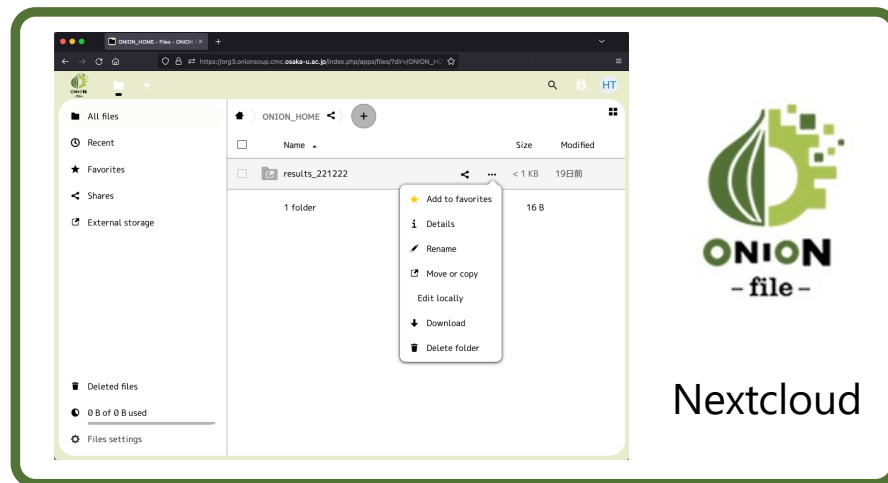


Recreated based on

https://ir.library.osaka-u.ac.jp/portal/about_en.html University of Osaka Open Science Symposium 2025

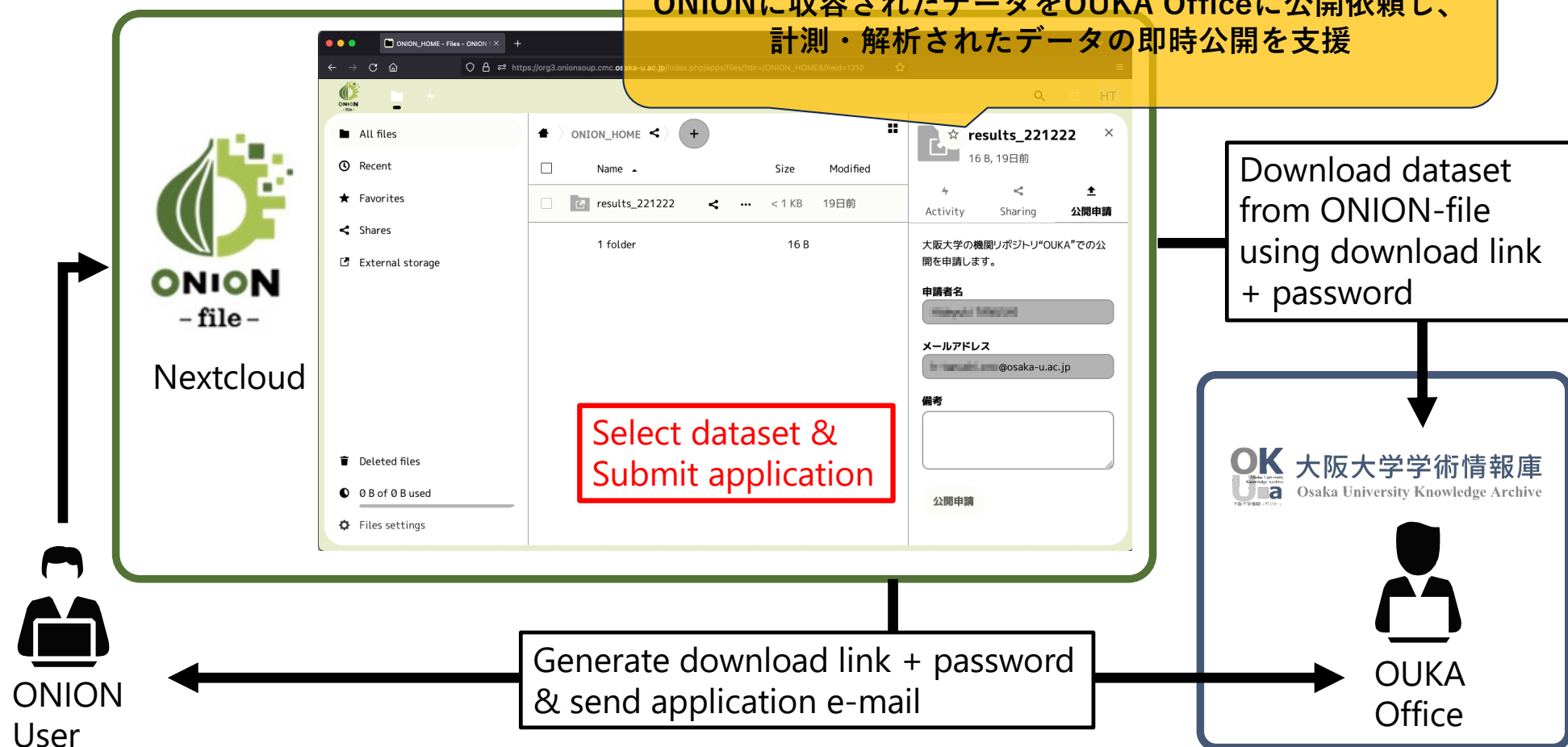
Courtesy of 田主英之(大阪大学 D3センター高性能計算・データ分析融合基盤協働研究所)

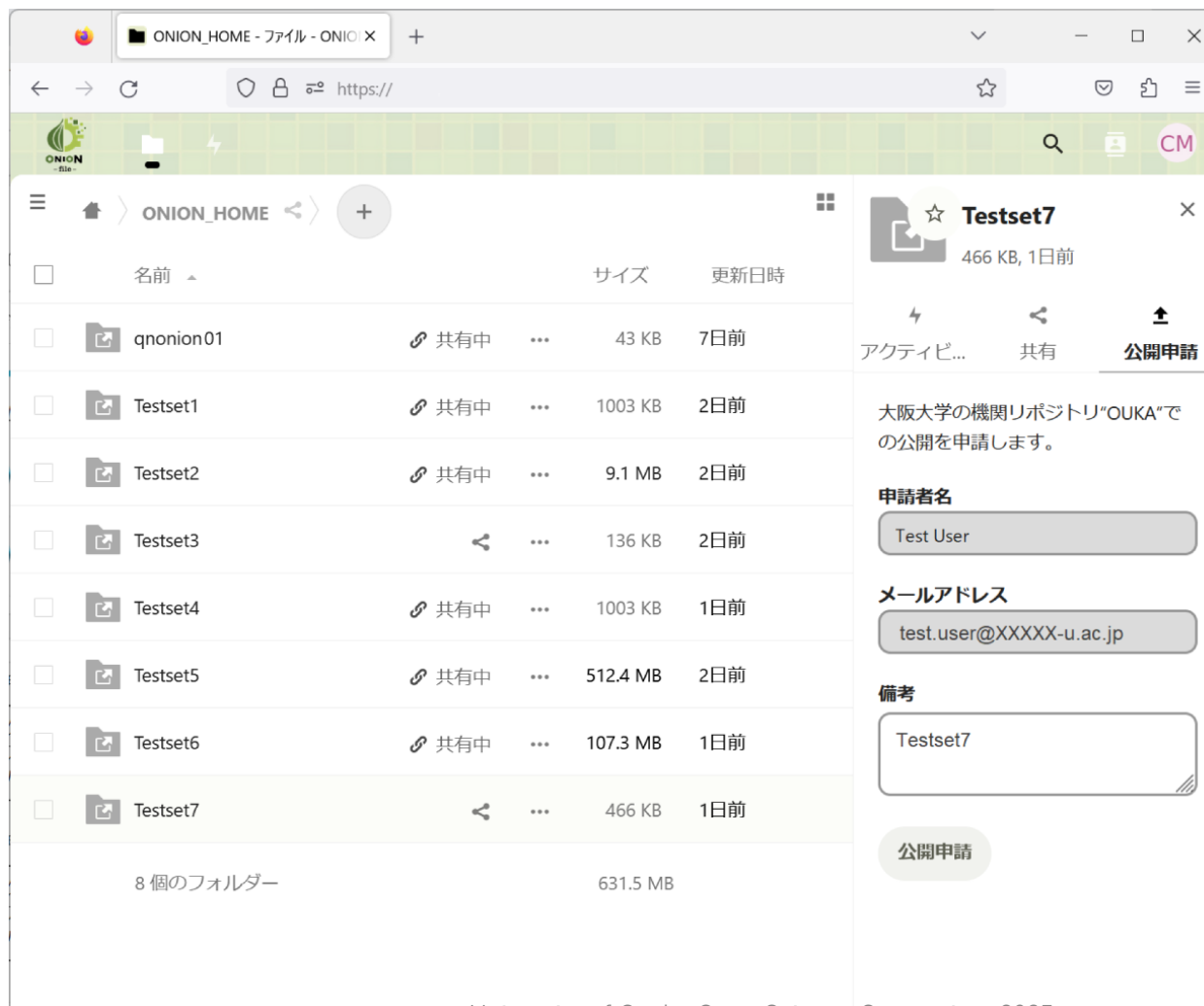
ONIONからOUKAに公開申請する際の問題



ONION-OUKAモジュール

ONIONに收容されたデータをOUKA Officeに公開依頼し、
計測・解析されたデータの即時公開を支援





The screenshot shows a web-based file management interface for 'ONION_HOME'. The main panel displays a list of folders with columns for name, size, and update time. The folder 'Testset7' is highlighted. A right-hand panel provides details for 'Testset7', including its size (466 KB) and update time (1 day ago). It also shows options for '公開申請' (Public Request) and a form to fill out the request, including fields for '申請者名' (Applicant Name) and 'メールアドレス' (Email Address).

名前	共有中	サイズ	更新日時
qnonion01	共有中	43 KB	7日前
Testset1	共有中	1003 KB	2日前
Testset2	共有中	9.1 MB	2日前
Testset3		136 KB	2日前
Testset4	共有中	1003 KB	1日前
Testset5	共有中	512.4 MB	2日前
Testset6	共有中	107.3 MB	1日前
Testset7		466 KB	1日前
8 個のフォルダー		631.5 MB	

Testset7
466 KB, 1日前

大阪大学の機関リポジトリ“OUKA”での公開を申請します。

申請者名
Test User

メールアドレス
test.user@XXXXX-u.ac.jp

備考
Testset7

公開申請



測定データ集約・配信システム

~科学計測機器からデータを収容する仕組み~



Courtesy of 古谷浩志 (大阪大学 コアファシリティ機構),

University of Osaka Open Science Symposium 2025

学内のデータ移動・集約需要を喚起

**全学
機器共用**

"部局"の研究リソースを"全学"でも共に活用する

大阪大の場合は、全部で300機器程度

約300機器・年間約8万件の利用

**これらの約300の共用機器が生み出すもの
→ 測定（研究）データ**

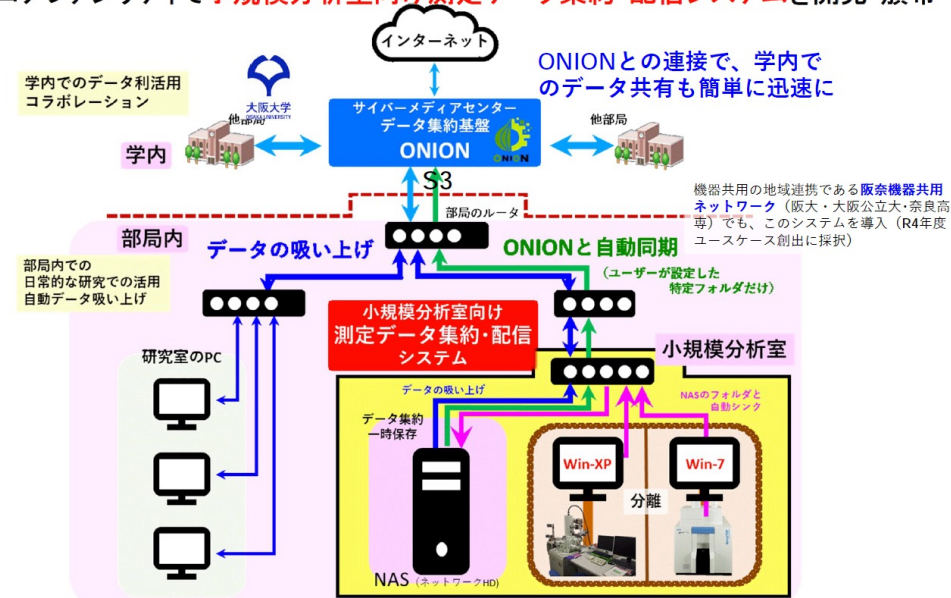


全体の取り纏め役（旗振り・調整役）がコアファシリティ



Courtesy of 古谷浩志 (大阪大学 コアファシリティ機構), "大阪大学コアファシリティ機構における測定データ流通・集約基盤の構築と阪奈機器共用ネットワークへの展開", 第1回北陸地区学術データ基盤セミナー ～コアファシリティ連携から研究データエコシステム構築を目指して～, 金沢, 2024年2月.

**ネットワーク経由での(部局内での)測定データの流通を実現
コアファシリティで小規模分析室向け測定データ集約・配信システムを開発・頒布**



測定データ集約・配信システムの利用・導入状況(1)

DB CENTER

工学研究科 分析センター

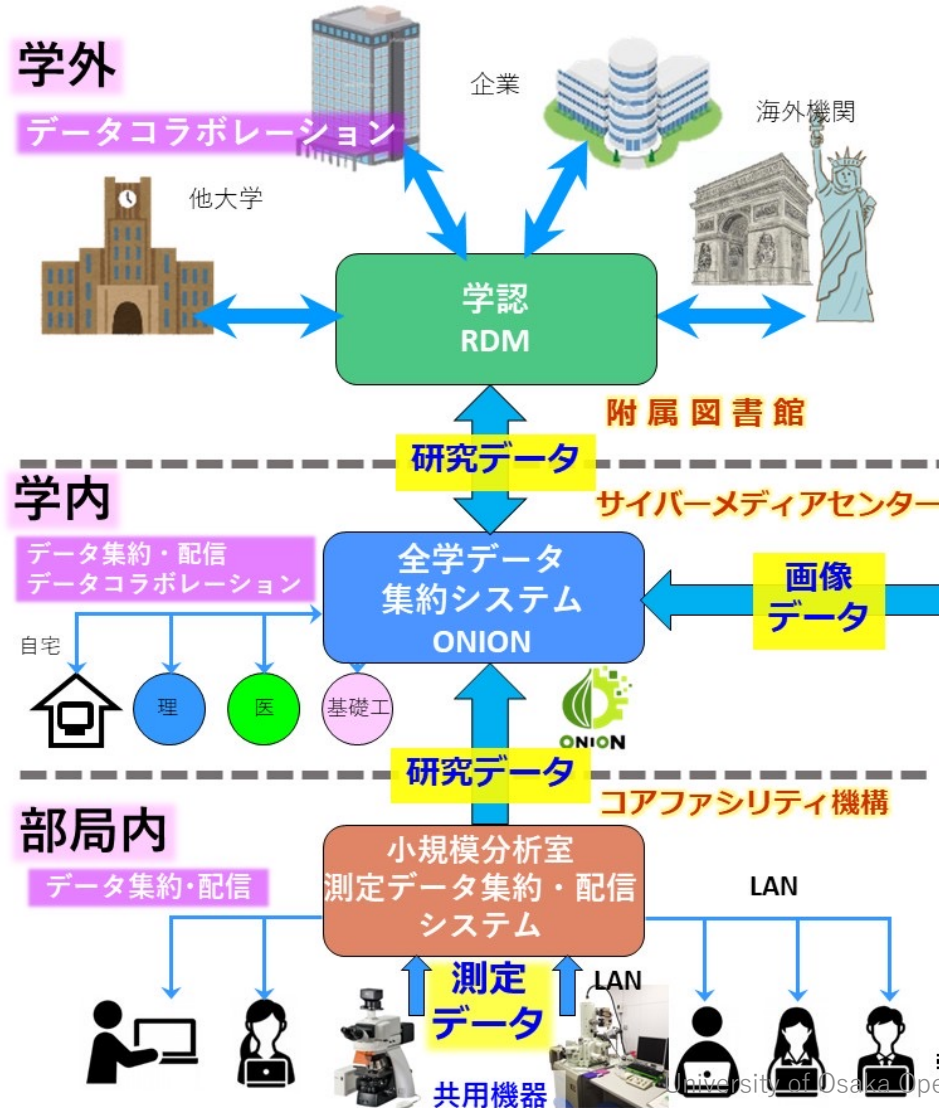
接続した共用分析機器：NRM装置 3台 + 質量分析装置 5台（合計8台） 機器担当者：2名

R3年度 機器の総利用件数：19070 件



R4年度から稼働する小規模分析室向け
測定データ集約配信システム

- R4年度からUSBメモリ使用を停止し、完全ネットワーク配信へ移行
 - R5年5月平日の平均転送ファイル数 340
- ⇒ 年間 約10万ファイルのペース



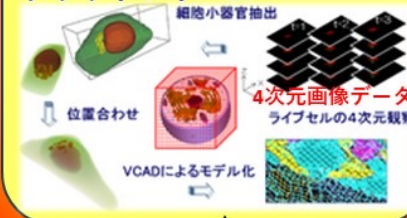
2025-2027年度 OUマスタープラン実現加速事業
「研究データの利活用を加速するmdx II /ONIONを
活用した全学画像解析クラウドプラットフォームの
構築と全学展開」
コアファシリティ機構、附属図書館、D3センター、
理化学研究所、WPI-PRIME

ONIONに研究データをアップすると
高価な解析ソフトの購入不要で
高機能解析ができる！

研究データ解析

mdx II

4D画像処理 VCAT5
プラットフォーム



利用・解析

さらなる発展にむけ

来歴管理システム

～計算過程を追跡・記録する来歴管理システム～



Courtesy of 並木悠太研究員大阪大学D3センター 高性能計算・データ分析融合基盤協働研究所),
University of Osaka Open Science Symposium 2025

OCTOPUS since Dec. 2017

OCTOPUS

Osaka university Cybermedia cenTer
Over-Petascale Universal Supercomputer



- ペタフロップス級ハイブリッド型スーパーコンピュータ

(Osaka university Cybermedia cenTer Over-Petascale Universal Supercomputer) OCTOPUS 1.46 PFlops

- 総理論演算性能 1.463 PFlops

2025年9月に
オープンサイエンスを促進する
計算・データプラットフォーム
として整備完了・サービス開始予定

CPU nodes: 236	
プロセッサ	Intel Xeon Gold 6126 (Skylake / 2.6 GHz 12コア) 2基
主記憶容量	192 GB
インターコネク	InfiniBand EDR (100 Gbps)
GPU nodes: 37	
プロセッサ	Intel Xeon Gold 6126 (Skylake / 2.6 GHz 12コア) 2基
主記憶容量	192 GB
インターコネク	NVIDIA Tesla P100 (NVLink) 4基 InfiniBand EDR (100 Gbps)
Many core nodes: 44	
プロセッサ	Intel Xeon Phi 7210 (Knights Landing / 1.3 GHz 64コア) 1基
主記憶容量	192 GB
インターコネク	InfiniBand EDR (100 Gbps)
Large memory nodes: 2	
プロセッサ	Intel Xeon Platinum 8153 (Skylake / 2.0 GHz 16コア) 8基
主記憶容量	6 TB
インターコネク	InfiniBand EDR (100 Gbps)
Storage	
ファイルシステム	DDN EXAScaler
容量	3.1 PB

次期OCTOPUSでの研究データの来歴管理システム/サービス

- デジタル表現されたデータを計算・解析処理する高性能計算基盤上で生成されるデータがどのように生成・作成されたかを追跡・記録

