

研究データの管理・利活用を支える メタデータ設計支援システムの開発と 計測データ管理への応用

もうじきおしまい

理化学研究所
大阪大学
(コアファシリティ機構)

小林 紀郎
古谷 浩志

オレが本家やで



祝 ノーベル生理学賞！

大阪大学「ワニ博士」



大阪大学
OSAKA UNIVERSITY

各部局で共用利用可能な研究設備・機器の例

これらを全学で利用可能とするのが**コアファシリティ機構**

(現在、コアファシリティが関与しているのは**約400機器**)

吹田キャンパス

超高圧電子顕微鏡センター

300万kVからクライオまで多数の電顕

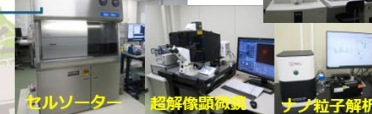
原子分解能TEM-EDS



Cryo-TEM (Titan-Krios) TEM-EDS 300万kV超高圧電顕

医学系研究科

日本のコアファシリティの原点



ニコンイメージングセンター



生命機能研究科

超解像顕微鏡+クライオ電顕
多数のイメージング装置



薬学研究科

有機化学・生体試料分析を中心とした装置群



CIDER

感染症研究・細胞解析などに関する最先端の51研究設備・機器！



工学研究科

時間分解レーザー分光装置などユニークな分析装置

400~600 MHz 溶液NMR



ナノ秒時間分解分光

MALDI

産業科学研究所

多数の最先端分析装置

TOF-SIMS (M6) 原子分解能TEM (JEM-ARM200F) 電子線回折構造解析装置 (SynergyED)



蛋白質研究所

日本最大級のNMR装置群



800 MHz 600 MHz 500 MHz

豊中キャンパス

2025年4月設置

質量分析センター (理学研究科+基礎工+コア)

質量分析に特化した拠点
先端材料・化学研究に必要な質量分析装置群を集約・整備中

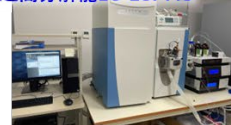
タンデム磁場型質量分析計



高分解能 UPLC-ESI-Q-TOF-MS



超高分解能LC-ESI-MS



SpiralTOF MALDI



Tandem-TOF MALDI

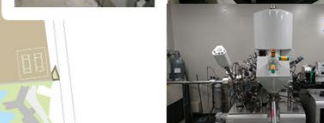


GC-QMS



基礎工学研究科

有機・無機分析を中心とした装置群



理学研究科

ひとつの大学に匹敵する程の多数かつ多様な分析装置群



分子間相互作用解析



大阪大学
OSAKA UNIVERSITY

約400機器
年間8万件の利用



共用機器



共用機器



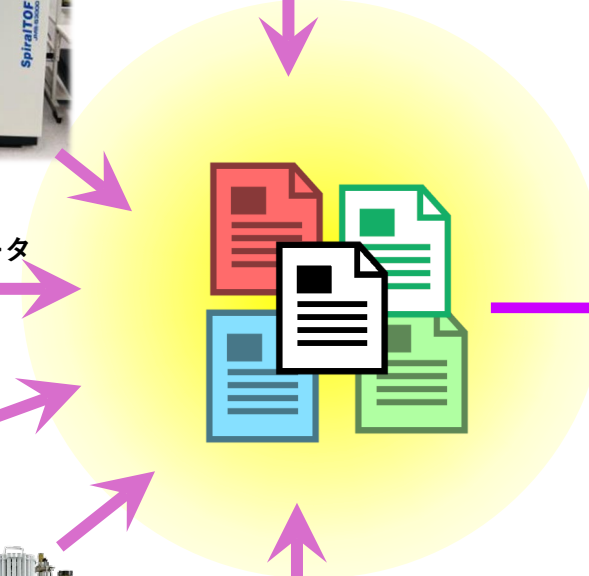
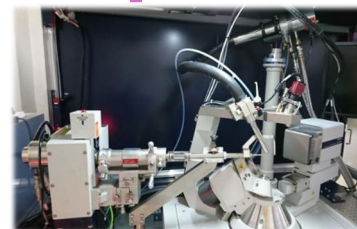
分析・測定データ



分析・測定データ



分析・測定データ



共用機器の利用で生まれる
分析・測定データは全て

**実験系
研究データ**

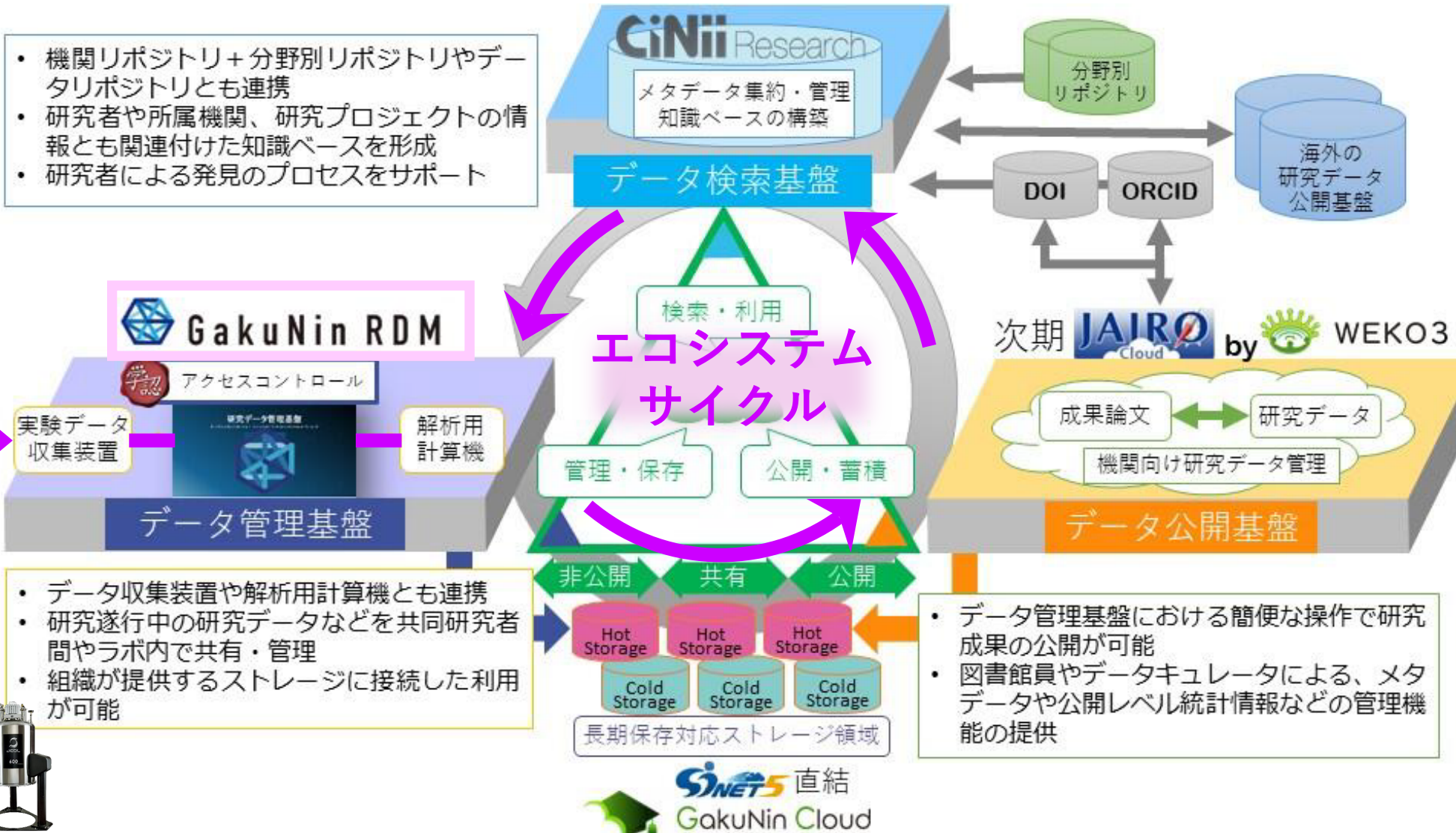
コアファシリティ

- 多くの理工生命系研究者が関わり
- 利用件数も非常に多く
- 限られた種類のデータフォーマット

“実験系”の研究データエコシステム
として非常に良いモデルケース

測定データと研究データエコシステム

このエコシステム・サイクルに、測定（研究）データを入れ込みたい！



測定データと研究データエコシステム

このエコシステム・サイクルに、測定（研究）データを入れ込みたい！

それ必要なのは

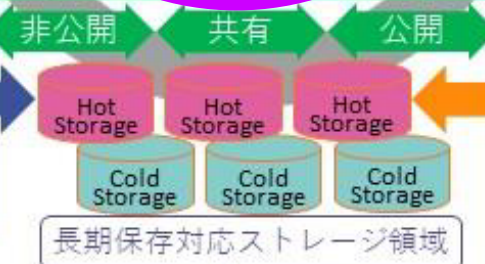
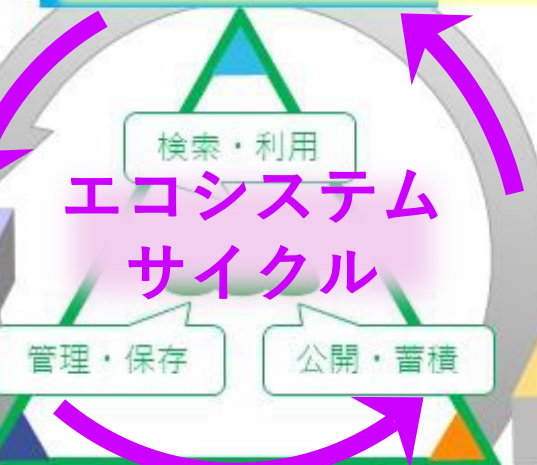
1. 測定データをITネットワーク上に持ち込むこと
2. 利活用に必須のメタデータを付帯させること

- ・ 機関リポジトリ+分野別リポジトリやデータリポジトリとも連携
- ・ 研究者や所属機関、研究プロジェクトの情報とも関連付けた知識ベースを形成
- ・ 研究者による発見のプロセスをサポート

測定
データ

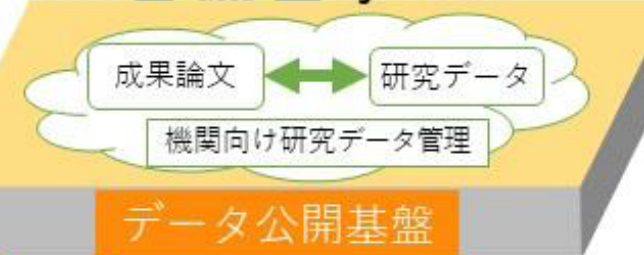


- ・ データ収集装置や解析用計算機とも連携
- ・ 研究遂行中の研究データなどを共同研究者間やラボ内で共有・管理
- ・ 組織が提供するストレージに接続した利用が可能



- ・ データ管理基盤における簡便な操作で研究成果の公開が可能
- ・ 図書館員やデータキュレータによる、メタデータや公開レベル統計情報などの管理機能の提供

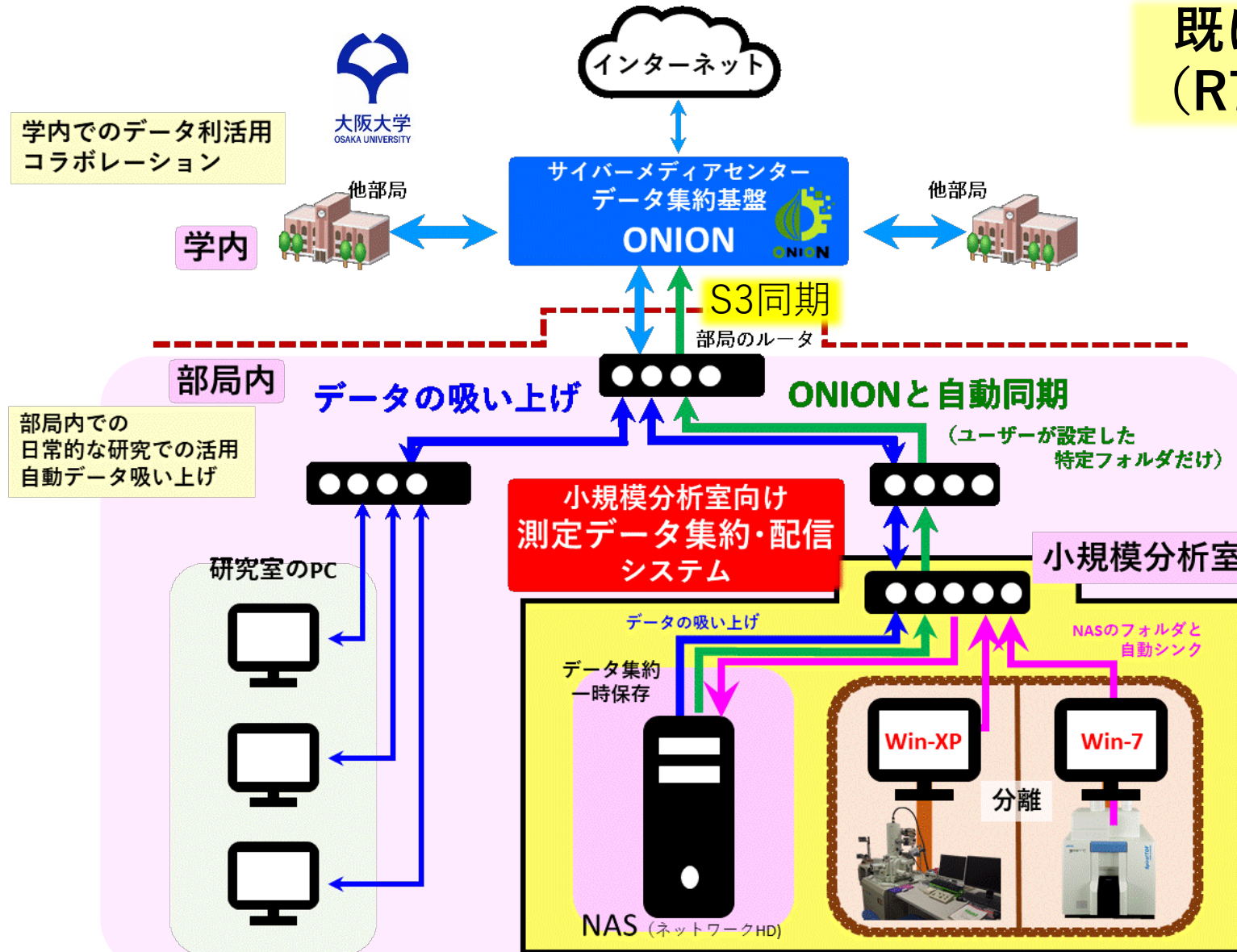
次期 JAIRO Cloud by WEKO3



測定データをネットワークに載せる

→ 小規模分析室向け測定データ集約・配信システム

既に11カ所に導入済み
(R7年度中に14カ所に)



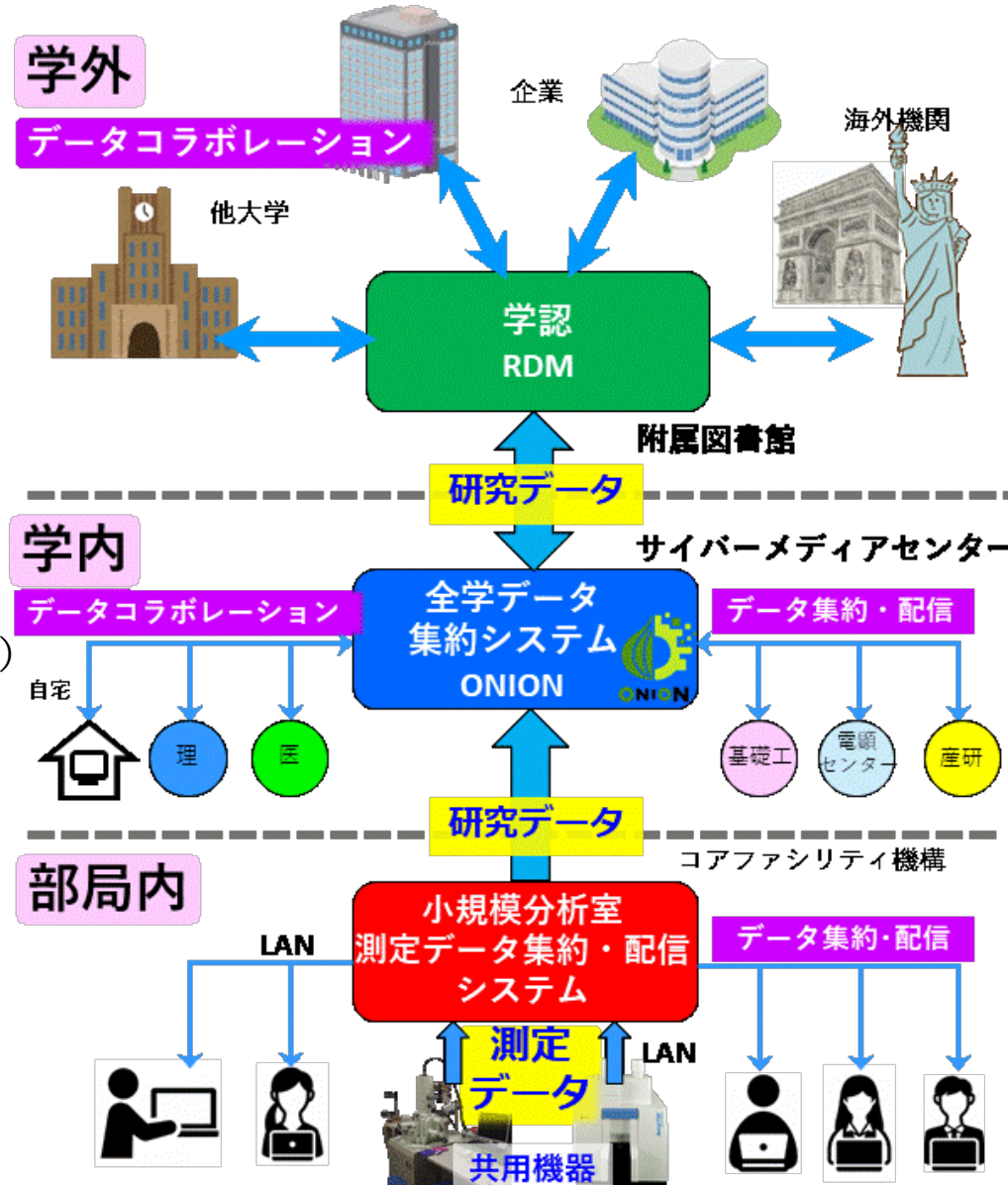
共用機器からの測定データ ネットワーク経由で一気通貫に流通・利活用する基盤は出来上がった！



大阪大学
OSAKA UNIVERSITY

附属図書館

オープン
データ推進



D3センター
(旧サイバーメディアセンター)
システム

コアファシリティ
データ
生産現場

測定データはネットワーク上
に乗るようになった！



次は利活用



利活用には
メタデータが必要！

メタデータのない
実験データは誰も使わない！

どうメタデータを付与・管理するか？

メタデータの付与はいつも面倒：
研究者にとって余分な仕事
どう付けて、**負担をどう減らすか？**

阪大のコンセプト

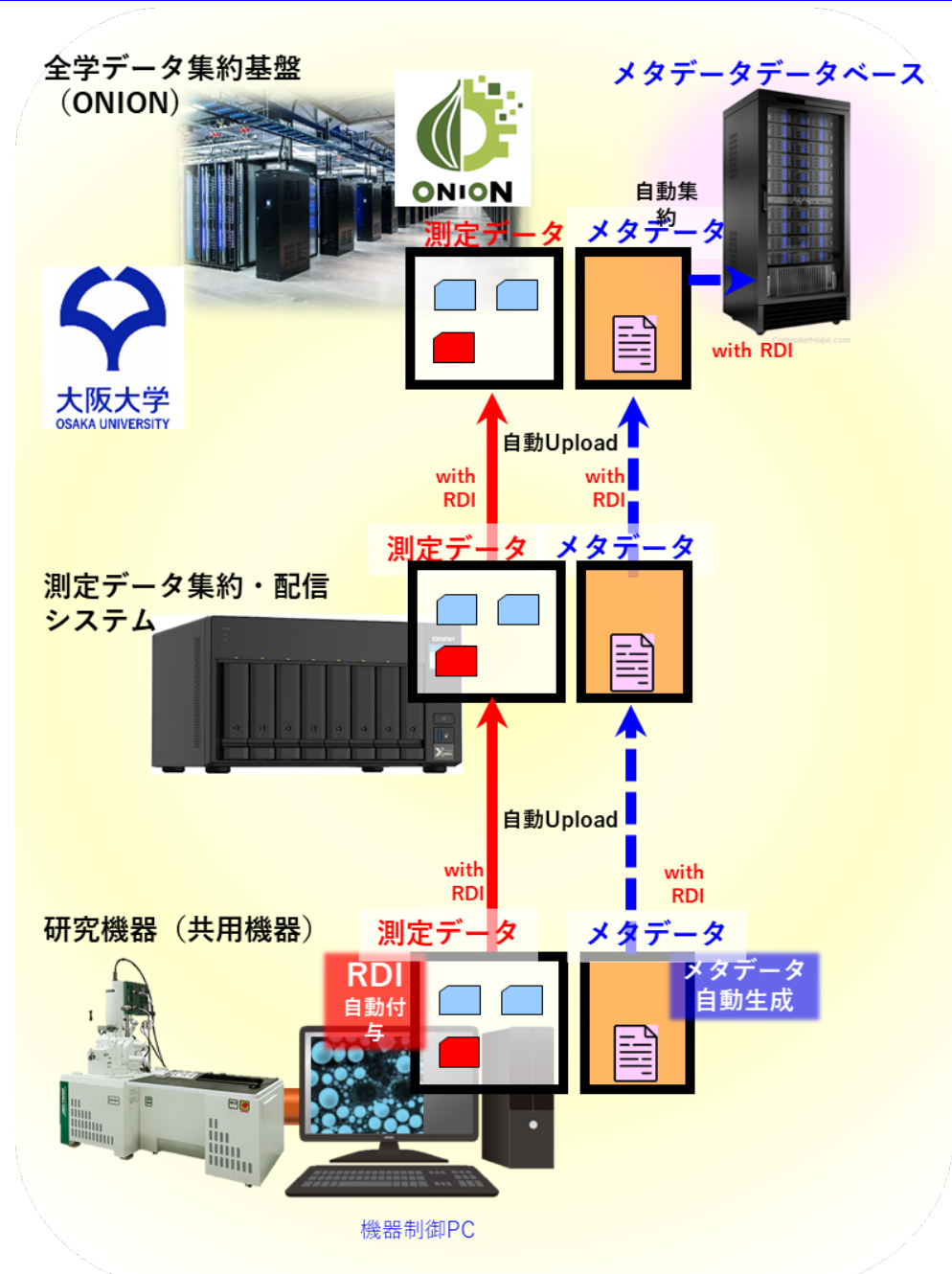
1. 測定データのネットワーク流通の流れと同じ経路でメタデータも流通させる
2. 研究機器から測定データをネットワーク転送する際に（**最低限の**）メタデータも作り、転送する
3. その際、測定データには管理用IDを付与
4. 全学データ集約基盤と同じレベルに、メタデータを転送・集約し管理する

H. Tanushi, H. Furutani, T. Hosomi, N. Kai, K. Harumoto, & S. Date. (2024). Towards Development of University-wide Data Aggregation and Management Infrastructure for Research Data Utilization. In 2024 IEEE 20th International Conference on e-Science (e-Science) (pp. 1–4). <https://doi.org/10.1109/e-Science62913.2024.10678692>

田主 英之, 春本 要, 伊達 進, 古谷 浩志, 甲斐 尚人, 大阪大学データビリティフロンティア機構 2023年度学際共創プロジェクト「OU 研究データマネジメントエコシステム構築を目指した実験系研究データメタデータ管理の検証モデルシステム構築」報告書, 2024年5月.

現在、共用NMRを対象としてコンセプト実証システムを構築中
同時に、共用NMRデータ用の**メタデータ**も検討・構築中

理研のメタデータ
設計ツールを活用



理研 生命科学実験研究メタデータ

理化学研究所 小林先生のグループが「生命科学実験研究データ」に関して調査し一般化
「生命科学実験研究データ」だけでなく、**一般的な「実験系研究データのメタデータ」に適用可能**

プロジェクト

- プロジェクトID
- タイトル
- 説明
- 作成者
- 作成者
- 研究責任者
- 連絡窓口
- 参考文献
- 実験

実験

- 実験ID
- タイトル
- 説明
- 測定の種類
- 技術の種類
- 測定基盤
- 参考文献
- 実験日
- 実験者
- 実験設計
- 測定
- データ解析

測定

- サンプル
- 測定条件
- データセット
- 計測日時
- 説明

データセット

- フォルダ
- ファイル
- 圧縮ファイル

測定条件

- 機器
- 機器生成メタデータファイル

機器

- 機器ID
- 機器名
- 所在地
- 管理者

データ解析

- 前処理
- 統計的データ解析
- 単変量解析
- 多変量解析
- アノテーション方法
- 可視化
- 解析結果データセット

サンプル

- サンプルID
- 名称
- 説明
- 化合物
- 生物サンプル
- サンプル調整

サンプル調整

- 処理法
- サンプル量
- サンプル抽出プロトコル
- 抽出法
- 標準化合物

60クラス、152項目を定義 青色部分は研究ごとに定義が必要

理研 生命科学実験研究メタデータ

理化学研究所 小林先生のグループが「生命科学実験研究データ」に関して調査し一般化
「生命科学実験研究データ」だけでなく、**一般的な「実験系研究データのメタデータ」に適用可能**

プロジェクト

- プロジェクトID
- タイトル
- 説明
- 作成者
- 作成者
- 研究責任者
- 連絡窓口
- 参考文献
- 実験

実験

- 実験ID
- タイトル
- 説明
- 測定の種類
- 技術の種類
- 測定基盤
- 参考文献
- 実験日
- 実験者
- 実験設計
- 測定
- データ解析

測定

- サンプル
- 測定条件
- データセット
- 計測日時
- 説明 ・ 測定者

共用
必須

データセット

- フォルダ
- ファイル
- 圧縮ファイル

一部
必要

測定条件

- 機器
- 機器生成メタデータファイル

共用
必須

機器

- 機器ID
- 機器名
- 所在地
- 管理者

共用
必須

データ解析

- 前処理
- 統計的データ解析
- 単変量解析
- 多変量解析
- アノテーション方法
- 可視化
- 解析結果データセット

サンプル

- サンプルID
- 名称
- 説明
- 化合物
- 生物サンプル
- サンプル調整

一部
必要

サンプル調整

- 処理法
- サンプル量
- サンプル抽出プロトコル
- 抽出法
- 標準化合物

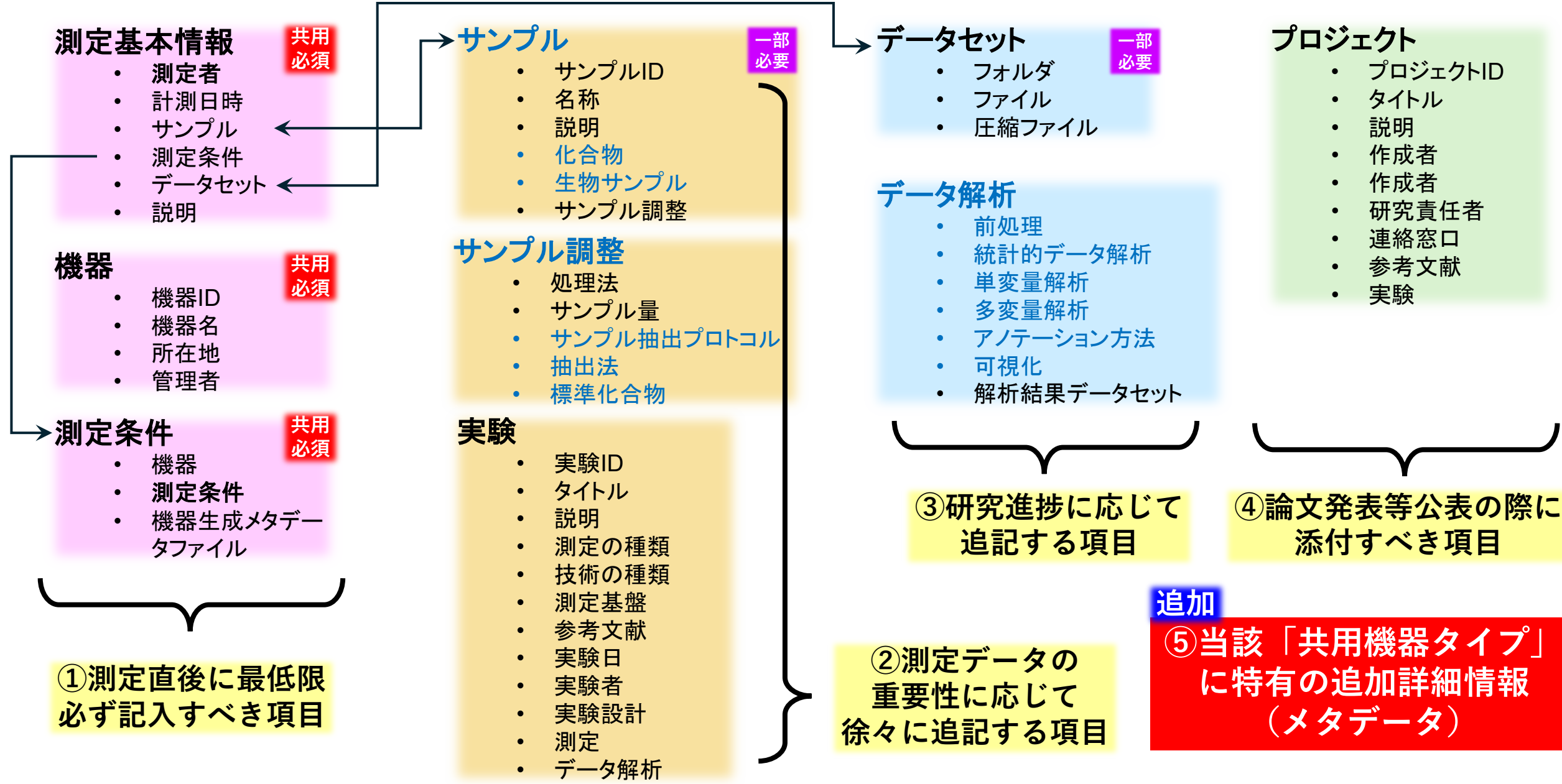
この概念を
共用機器測定データ
のメタデータの
一般化に活用

60クラス、152項目を定義 青色部分は研究ごとに定義が必要

理研 生命科学実験研究メタデータを「共用機器測定メタデータ用」に整理

大項目とClass

一部項目を追加・修正



共通化された「共用機器測定用のメタデータ形式」のデザイン

測定基本情報

- 測定者
- 計測日時
- サンプル
- 測定条件
- データセット
- 説明

共用
必須

機器

- 機器ID
- 機器名
- 所在地
- 管理者

共用
必須

測定条件

- 機器
- 測定条件
- 機器生成メタデータファイル

共用
必須

サンプル

- サンプルID
- 名称
- 説明
- 化合物
- 生物サンプル
- サンプル調整

一部
必要

サンプル調整

- 処理法
- サンプル量
- サンプル抽出プロトコル
- 抽出法
- 標準化合物

実験

- 実験ID
- タイトル
- 説明
- 測定の種類
- 技術の種類
- 測定基盤
- 参考文献
- 実験日
- 実験者
- 実験設計
- 測定
- データ解析

データセット

- フォルダ
- ファイル
- 圧縮ファイル

一部
必要

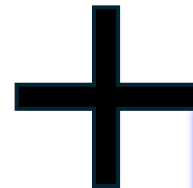
データ解析

- 前処理
- 統計的データ解析
- 単変量解析
- 多変量解析
- アノテーション方法
- 可視化
- 解析結果データセット

プロジェクト

- プロジェクトID
- タイトル
- 説明
- 作成者
- 研究責任者
- 連絡窓口
- 参考文献
- 実験

共通項目



当該共用機器
タイプに特有の項目

- オプションモジュール
- 説明
- 機器の利用形態
- 共用システム名
- 機器利用方式
- 機器詳細情報
- 参考文献
- 実験

共用機器タイプ

NMR・電子顕微鏡・質量分析
X線構造解析・表面分析・分光分析
など

各部分を
理研メタデータ設計支援ツール
で設計・管理

Meta-data Toolkit



① ユーザ管理 → ② プロジェクト
作成・管理 → ③ スキーマ
編集・管理・承認・公開



- 同ツール上で各共用機器ごとに設計
- 共通部分は使い回し
- 多人数で協力してデザイン

現在、NMRを対象にデザインを
開始したところ