

海外調査進捗報告

国立情報学研究所

山地一禎

2019年5月16日

オープンサイエンスの深化と推進に関する検討委員会
(第24期・第5回)

宿題

データ駆動型（含：研究データ管理）の研究をサポートする機関のサービスについて、

- 海外機関におけるサービス展開の例を具体的に調べる。
- そのサービスにより研究促進に繋がった例を具体的に調べる。

ことにより、日本におけるオープンサイエンスの深化と推進に、何が必要かを考える材料をそろえる。

方法

- ヨーロッパとオーストラリアの先行機関等を訪問し、サービスの詳細や成果についてインタビュー調査を実施
 - 時間の都合上、直接訪問できない北米については、これまでの交流を生かして、TV会議等で情報収集
- ヨーロッパ
 - 担当：京大 青木学聡、NII 藤原一毅
 - 日程：2019年6月16日～22日
- オーストラリア
 - 担当：NII 常川真央、NII 込山悠介
 - 日程：2019年5月13日～16日

3

オーストラリア編

常川真央、込山悠介

4

背景と目的

- 背景：

- オーストラリアでは研究者だけでなくHDR(High Degree Research) Candidate に対してもDMPの提出が義務付け
- RDM環境の整備とトレーニング支援に各大学が取り組んでいる

- 目的：

- DMPを軸とする先進的なRDMサービスやトレーニングに取り組む大学にインタビュー調査
- 日本国内に必要なRDMサービスの機能やトレーニングプログラムの要件を抽出・分析

5

調査対象

1. The University of Queensland (UQ)

- 規模：Academic staff 6,613, Students 52,331 (2018) 豪州8大学
- 観点：RDM中のmaDMPから、動的にストレージを確保できるシステムが実運用段階に有る

2. University of New South Wales (UNSW)

- 規模：Academic Staff 2,838; Professional Staff 3,390; (2018) 1)
Students 59,781 (2017) 1)豪州8大学
- 観点：ポリシーを軸にしたRDMツールと教材の連携

3. University of Sydney (USYD)

- 規模：Administrative Staff 5,350 (2015); Students 59,129(2017) 2)
豪州8大学
- 観点：データサイエンスのトレーニングプログラムと環境が充実

1) <https://www.unsw.edu.au/sites/default/files/documents/2018-At-a-Glance.pdf>

2) <https://sydney.edu.au/content/dam/corporate/documents/about-us/values-and-visions/Annual-report-2.pdf>

6

UQのDMPを軸にした動的な統合研究基盤 (1/2)

- UQ RDM : **研究データ管理サービス** 1)
 - Data Management Record (DMR)というアプローチを採用
 - 研究の計画やステータス、FAIR原則に基づくデータ管理情報を内包する常時更新可能なレコードを研究者が作成
 - 自動的なDMP生成やストレージ確保を実現
- eSpace : **機関リポジトリ** 2) 3)
 - UQではeSpaceを研究者のショーケースとして利用 (リポジトリでマッチメイキング)
 - Institutional Research 機能が有り、研究者の評価可
- MeDiCI : **データ収集・解析基盤** 4)
 - UQRDMのDMR機能と連携し仮想マシン、ストレージを確保
 - 実験装置からのデータ集約、HPC連携、コールドストレージでのデータアーカイブを自動化

1) <https://guides.library.uq.edu.au/for-researchers/uq-research-data-manager/about>

2) <https://espace.library.uq.edu.au/dashboard>

3) <https://guides.library.uq.edu.au/for-researchers/uqespace-publications-datasets/my-dashboard>

4) <https://rcc.uq.edu.au/data-storage>

UQのDMPを軸にした動的なRDM基盤を実現 (2/2)

- RDAとの関係性
 - UQRDMのDMRアプローチは、Research Data Alliance (RDA) のDMP Common Standards WG1)とは独立
- RDMおよびDMRに関するUQの課題
 - **eSpaceとの自動連携はこれから実装予定**
 - **UQRDMではDMPの入力項目の設計も試行錯誤中**
 - 学内のファンディングデータベースとUQRDMのシステム連携はまだ不十分 (他部局のデータベースへのリンクのみ)
- UQRDMのフロントエンドにて**ストレージ割当時のプロビジョニングは開発済**
 - 新規プロジェクトは5~20分で準備
- UQRDM中で助成事業を検索する機能
 - 研究者とFAのマッチング
 - 研究者の研究アイデアと研究助成を結びつける

1) <https://www.rd-alliance.org/groups/dmp-common-standards-wg>

UNSWのポリシーを軸にRDMツールと教材を連携

- ポリシー：整理された研究データ分類表
 - 各部局と協議しながら作成
 - データの性質によってカテゴリ分け
 - カテゴリごとにどのストレージを使うべきか明示
- トレーニング：RDMオンライン教材&ワークショップ
 - LMS(Smart Sparrow)を導入して独自にオンライン教材を作成
 - RDMの一般知識と併せて学内ツールのインストラクションも
 - 学習履歴の分析：離脱率や回答の流れを集計・可視化
 - Hacky Hour：RDMを含めた研究ツールのレクチャーとワークショップ
 - ResBaz：国際的なHacky Hourとしてのカンファレンス（研究者、大学院生のマッチングメイキングの役割もある）
- ツール：ResData
 - DMPとデータセットのメタデータ作成やストレージ自動確保

9

Sydney Informatics Hub, University of Sydney

- データサイエンスのトレーニング教材を提供
 - HPC(High Performance Computing)
 - データ加工の方法
 - 分野特化のデータ分析手法（例：ゲノム）
 - Python やRによる機械学習
 - 研究データ管理やストレージのインストラクション
 - Etc...
- SIHもHacky Hourを開催し、ResBazに参加

ヨーロッパ編

青木学聡、藤原一毅

11

概要

- 目的
 - 研究データ解析システムの先進事例を調査する
 - ユーザー（研究者）のサポート/トレーニング体制について事例を集める
- 調査対象
 1. スイス連邦工科大学 チューリッヒ校 (ETH Zurich)
 2. スイス連邦工科大学 ローザンヌ校 (EPFL)
 3. ゲオルク・アウグスト大学ゲッティンゲン (GWDG)

12

ETH Zurich: OpenBIS

- Laboratory Information Management System (LIMS) の一種
- 主にバイオ系のラボで利用
- 測定機器からの0次出力の収集と解析ワークフローを統合的に扱うシステム
- OpenBISの機能、開発プロセス、利用シーン、図書館等との協業によるサポートについて調査

13

ETH Lausanne: AiiDA

- 再現可能なシミュレーションのための計算サービスを含む統合インフラ
- 主に材料科学系のラボで利用
- 公開データベースとの連携、スーパーコンピュータでの解析、結果の共有を可能

14

GWDG

- 大学の情報基盤センターであると同時に、マックス・プランク研究所におけるITセンターの役割を担っている組織
- RDMのための各種インフラ開発に参画
- 研究者にITトレーニングコースとサポートデスクを提供