

提言作成のための 情報収集調査報告

2019年7月24日

青木学聡（京大）、藤原一毅（NII）
込山悠介（NII）、常川真央（NII）
まとめ：山地一禎（NII）

欧州調査

日程：2019年6月17日～20日

メンバー：青木学聡（京大）、藤原一毅（NII）

1. スイス連邦工科大学チューリッヒ校 (ETH Zürich)

- ・相手方: Caterina Barillari 氏 (Scientific IT Services) ら4名
- ・研究データ管理システム OpenBIS の開発・運営に関する情報交換、および同校図書館が運営するRDMサポート業務に関する情報交換

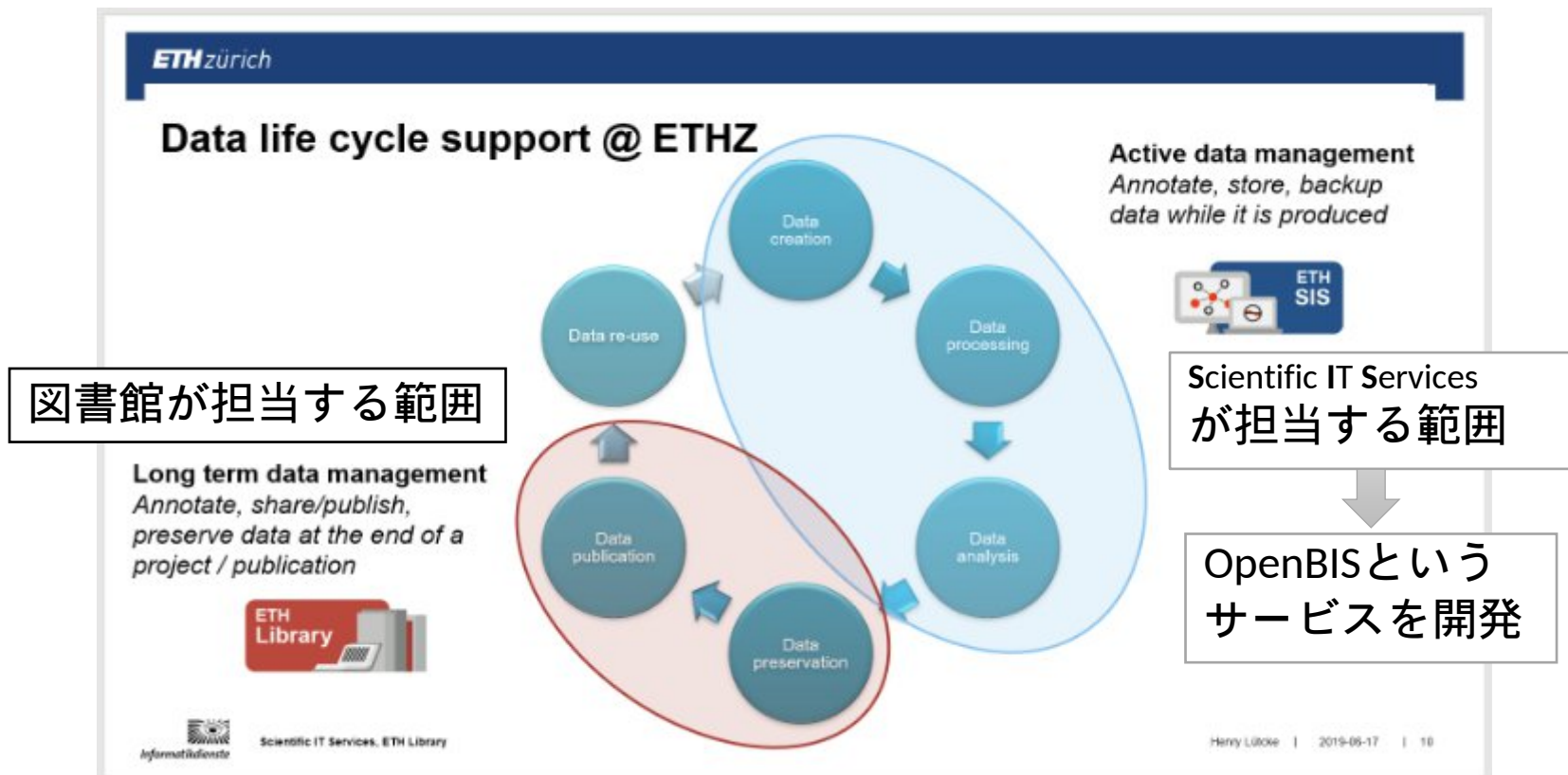
2. スイス連邦工科大学ローザンヌ校 (EPFL)

- ・相手方: Giovanni Pizzi 氏 (Laboratory of theory and simulation of materials) ら3名
- ・研究データ解析システム AiiDA の開発・運営に関する情報交換

3. ゲオルク・アウグスト大学ゲッティンゲン (GWDG)

- ・相手方: Ramin Yahyapour 教授 (Managing Director) ら3名
- ・同大学が運営している研究支援ICTサービスに関する情報交換

ETHZ : 図書館と基盤センターによるサポート



LibraryとSISが合同で、教育プログラムも実施

- RDM Summer School (5日間の集中コースにより単位認定)
- RDM Workshop

個別のツール、システムの利用方法についても、SIS、Libraryがそれぞれセミナーを 精力的に開催

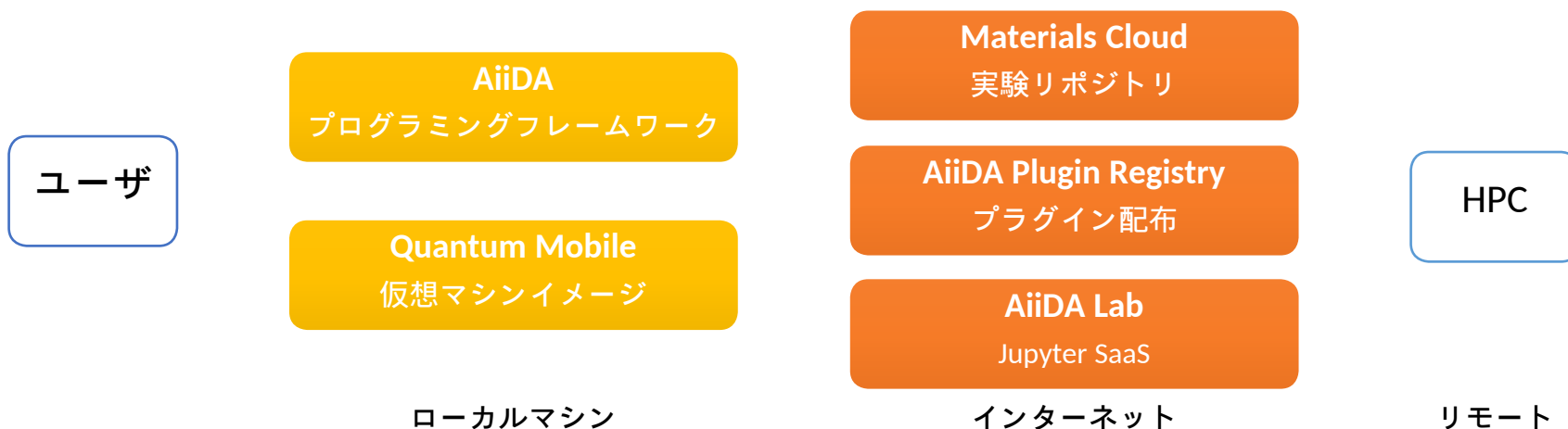
ETHZ : OpenBISの主な機能

研究室での統合的な情報管理を実現するために
以下の機能を提供（2019年現在）

- ・ インベントリ（ファイル、ドキュメント）サーバー
 - ・ 研究室のメンバー間でのデータ共有
- ・ 電子ラボノート（Electronic Laboratory Note, ELN）
 - ・ インベントリオブジェクトへのリンク 埋め込みなど、ラボのリソースと関係を明確に記述
 - ・ オブジェクトブラウザにより、オブジェクト間の依存関係、ワークフローを可視化
- ・ Jupyter Notebook等、オンサイト解析環境との連携
- ・ 外部ワークフローエンジンとの連携
- ・ 外部の大規模データストレージへの参照機能

EPFL : AiiDA

- ・シミュレーション作業を自動化 (automate) ・組織化 (organize) するとともに、データの来歴を保存し再現可能性を保証するためのツール (の必要を認識)
 - ・自動化されたワークフローに科学者の知識を埋め込む
 - ・データの来歴を自動的に追跡・保存する
 - ・FAIR原則にもとづき再現性が保証されたシミュレーションを共有する
- ・AiiDAは、EPFLで開発した上記をサポートするツール群



EPFL : AiiDA

シミュレーションに使ったデータとコードを、
AiiDAでそのまま実行可能な形で提供 (2018/2 ~)

materialscloud:2017.0008   

Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds

Authors: Nicolas Mounet^{1*}, Marco Gibertini¹, Philippe Schwaller¹, Davide Campi¹, Andrius Merkys^{1,2}, Antimo Marrazzo¹, Thibault Sohier¹, Ivano E. Castelli¹, Andrea Cepellotti¹, Giovanni Pizzi¹, Nicola Marzari^{1*}

1 Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
2 Vilnius University Institute of Biotechnology, Sauletekio al. 7, LT-10257 Vilnius, Lithuania

* Corresponding authors emails: nicolas.mounet@epfl.ch, nicola.marzari@epfl.ch

DOI [10.24435/materialscloud:2017.0008/v2](https://doi.org/10.24435/materialscloud:2017.0008/v2) (version v2, submitted on 21 March 2018)

How to cite this entry
Nicolas Mounet, Marco Gibertini, Philippe Schwaller, Davide Campi, Andrius Merkys, Antimo Marrazzo, Thibault Sohier, Ivano E. Castelli, Andrea Cepellotti, Giovanni Pizzi, Nicola Marzari, *Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds*, Materials Cloud Archive [2018], doi: [10.24435/materialscloud:2017.0008/v2](https://doi.org/10.24435/materialscloud:2017.0008/v2).

Description
Two-dimensional (2D) materials have emerged as promising candidates for next-generation electronic and optoelectronic applications. Yet, only a few dozens of 2D materials have been successfully synthesized or exfoliated. Here, we search for novel 2D materials that can be easily exfoliated from their parent compounds. Starting from 108423 unique, experimentally known three-dimensional compounds we identify a subset of 5619 that appear layered according to robust geometric and bonding criteria. High-throughput calculations using van-der-Waals density-functional theory, validated against experimental structural data and calculated random-phase-approximation binding energies, allow to identify 1825 compounds that are either easily or potentially exfoliable. In particular, the subset of 1036 easily exfoliable cases provides novel structural prototypes and simple ternary compounds as well as a large portfolio of materials to search from for optimal properties. For a subset of 258 compounds we explore vibrational, electronic, magnetic, and topological properties, identifying 56 ferromagnetic and antiferromagnetic systems, including half-metals and half-semiconductors. This archive entry contains the database of 2D materials (structural parameters, band structures, binding energies, etc.) together with the provenance of all data and calculations as stored by AiiDA.

Materials Cloud sections using this data

- Select 2d materials via interactive periodic table and view their properties (with links to provenance)
- Explore interface providing access to the full database

DOIs assigned → [10.24435/materialscloud:2017.0008/v2](https://doi.org/10.24435/materialscloud:2017.0008/v2)

Direct links to Discover & Explore → [Select 2d materials via interactive periodic table and view their properties \(with links to provenance\)](#)
[Explore interface providing access to the full database](#)

FAIRsharing.org
re3data.org

+

Recommended data repository by Nature's journal Scientific Data

GWDG : 汎用RDM基盤

- Gessellschaft für Datenverarbeitung mbH Göttingen (GWDG)は、ゲッチンゲン大学とマックス-プランク研究所に対し、ITサービス全般を提供するNPO
 - 提供サービスは、電子メール、認証システム、各種ホスティングサービス、ネットワークサービス、HPCのオペレーション等多岐
 - RDMに関するツール群も、このサービスの一環として提供
- 汎用RDM基盤
 - データマネジメントの要素技術を「Access(ユーザーインターフェース)」「Service(データ処理)」「Data(ストレージ)」の階層に分け、体系的な開発を実施

SCIENTIFIC INFORMATION INFRASTRUCTURE ON THE POLITICAL AGENDA

"We will develop a strategy for the digital transformation in science, i.e. for enhanced accessibility of research data. Together with the federal states we will fund a Council for Scientific Information Infrastructures which coordinates stakeholders in the national science system and drives the development of strategies and standards across disciplines and institutions."

Coalition agreement CDU, CSU and SPD
18. Legislation period, December 2013

TO FACILITATE AND ACCELERATE DIGITAL TRANSFORMATION IN THE SCIENCES ON A NATIONAL LEVEL

The Council will monitor transitions in the German science system at large and give practical recommendations to Science and Government. It will specifically

- > Provide foresight on the development of digital science
- > Facilitate coordination of existing activities
- > Identify potential synergies between the multitude of actors and new fields of action
- > Facilitate cooperation within the science system
- > Monitor transition on an international scale



The screenshot shows the GWDG CDSTAR website. The header includes navigation links: NEWS, SERVICES, RESEARCH & EDUCATION, ABOUT US, SUPPORT, and MY PORTAL. The main content area features the GWDG CDSTAR logo and a description of the system. A sidebar on the right contains a 'Table of Contents' with links to various resources. The footer mentions the Apache2 license and the availability of the software as an individual service or product.

GWDG : Göttingen eResearch Alliance (eRA)

・役割

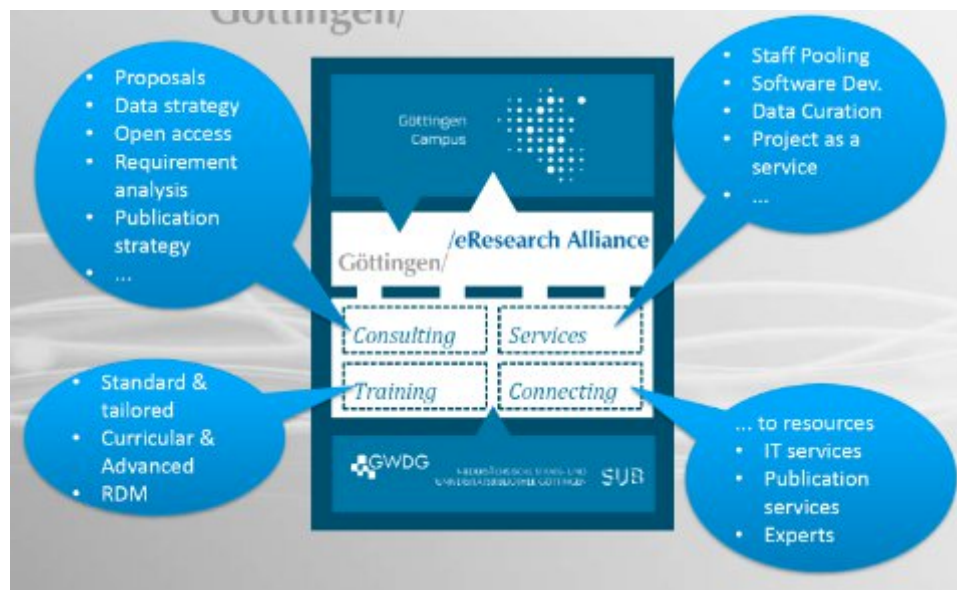
- ・ GWDGとSUB (Göttingen State and University Library) の専門家と共に、RDMを含むeResearch 全般に関わるトータルサポートと教育を実施

・支援

- ・ 大型プロジェクト申請時には、Research Officeと協力し、DMP作成支援と、eScience インフラ構築に関するコンサルティングを実施
- ・ RDMソリューションの持続的な適切なコストモデルを検討するプロジェクトも実施

・トレーニング

- ・ データサイエンス教育では、2017年より2週間 4 ECTSのサマースクールを開講



豪州調査

日程：2019年6月17日～20日

メンバー：込山悠介（NII）、常川真央（NII）

1. The University of Queensland (UQ)

- ・相手方: Joe Shapter教授(Pro Vice Chancellor Research Infrastructure)ら4名
- ・研究データ管理システム UQ RDMの開発・運営に関する情報交換

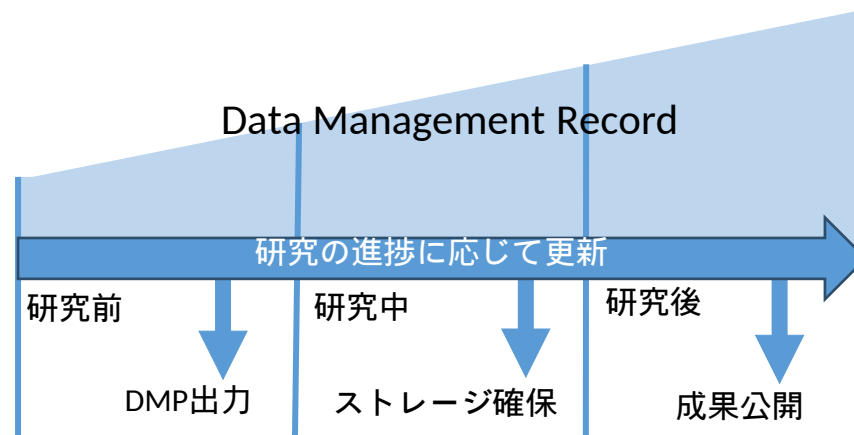
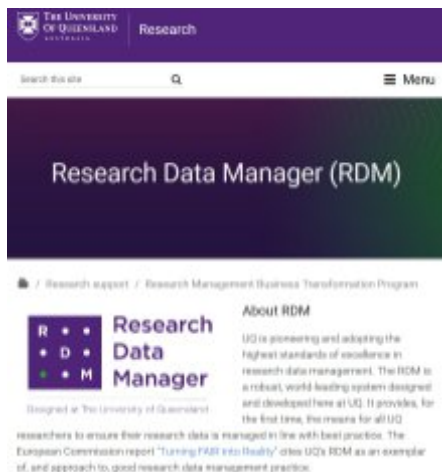
2. University of New South Wales (UNSW)

- ・相手方: Maude Frances氏(Associate Director, Scholarly Communications and Repositories)ら4名
- ・同大学における研究データポリシー、研究データ管理システム、RDMトレーニング教材作成の取り組みに関する情報交換

3. University of Sydney (USYD)

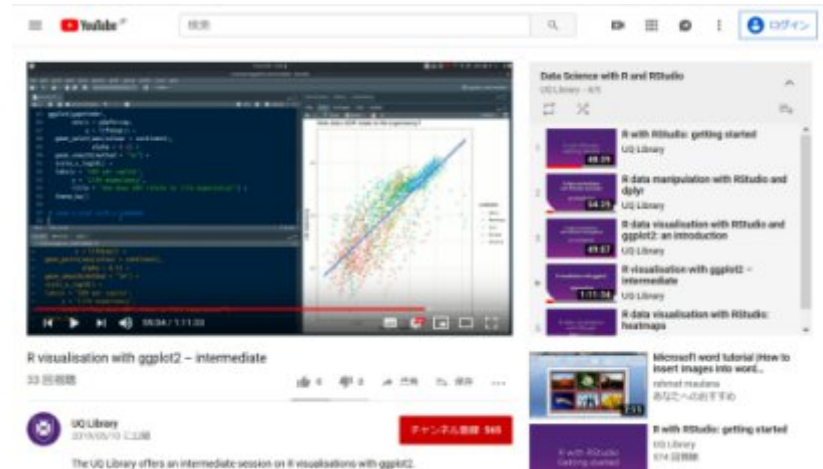
- ・相手方: Andrew Janke氏(Associate Director, Research Technology)ら4名
- ・同大学が提供するデータサイエンスのトレーニングプログラム Sydney Informatics Hub の運営に関する情報交換

UQ : UQ Research Data Manager



- RDMプラットフォームの中で、研究プロジェクトの開始から終了までの一連のステータスを1つの「レコード」として作成・更新
- 研究の過程で必要に応じて情報を利活用
 - 研究前：FAとのマッチング、DMPの出力
 - 研究中：ストレージの確保、研究データのメタデータ作成
 - 研究後：報告書の提出、データ保管、リポジトリへのデータ公開

UQ：図書館によるデータリテラシー教育



- Research Data Management Best Practice with UQ RDM 1)
 - UQ RDMを使用したRDMトレーニングコースを2箇 月に1回開催
- トレーニングワークショップの 実施
 - 電子ラボノートの利用
 - Data Science with R and Rstudio 2)
- UQRDM Pop-Up Sessions (UQ Hacky Hour)
 - 毎週決まった時間に開催されるセッション

1) <https://staffdevelopment.hr.uq.edu.au/course/RRPRDM>

2) https://gitlab.com/stragu/DSH/blob/master/R/rstudio_intro/rstudio_intro.md

UNSW : ResData

Storing your Research Data

Storage Options	UNSW On-campus & Tiers	Data Archive	Private Cloud	Public Data
Storage Type	On-premise	On-premise	On-premise	On-premise
Subject Class Classification	On-premise	On-premise	On-premise	On-premise
Researcher availability	On-premise	On-premise	On-premise	On-premise
Storage Link	On-premise	On-premise	On-premise	On-premise
Version Control	On-premise	On-premise	On-premise	On-premise
Researcher Time	On-premise	On-premise	On-premise	On-premise
Project Data	On-premise	On-premise	On-premise	On-premise

ResData

Welcome Elizabeth Stokes (Library Repository Services, University Library)

Databases

Title	Updated	Status	Action
Taxonomy of entomological disease	20-01-2018	Pending	Manage
Muscle protein synthesis and nutrition	20-02-2018	Pending (DOI)	Manage
Assessing microtubule status in the presence of information	20-02-2018	Draft	Manage
Intermittent deficiency in adolescents	20-02-2018	Draft	Manage

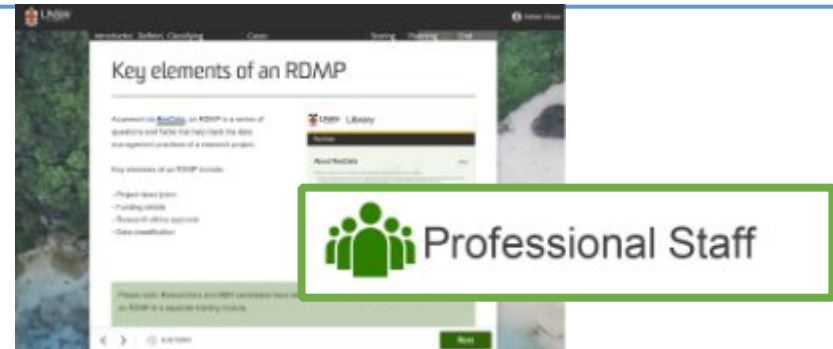
Project/grants

You are participating in the following project/grant(s):

Name	Funding Scheme	Role	Action
Role of microtubule in treating inflammatory breast disease	ARC	Contributor	Manage

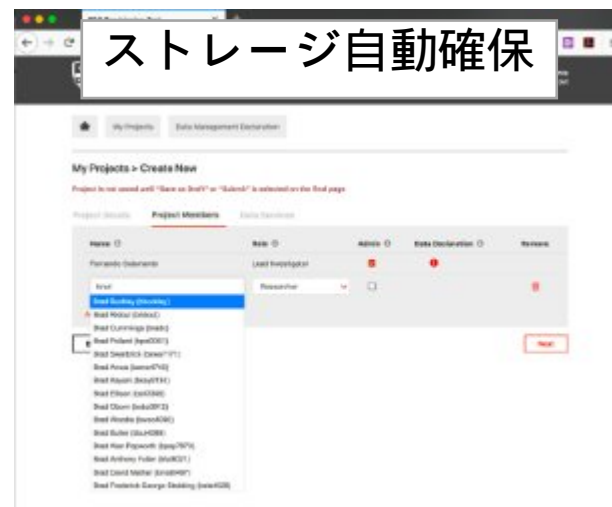
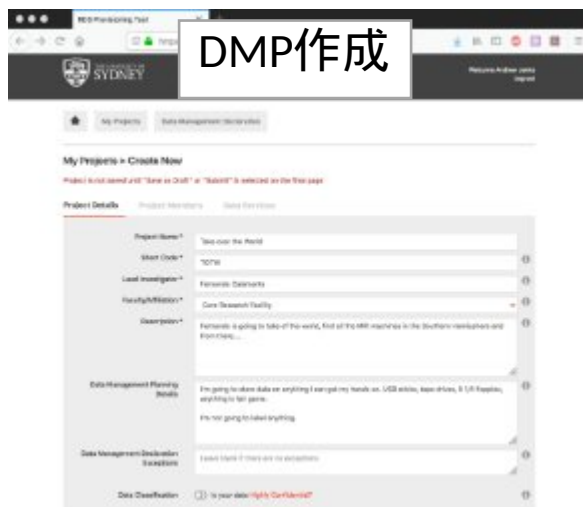
- 判別表 “Storing your Research Data” を用いて研究者自身がデータ性質を格付け
- データの種類ごとに適切なストレージをシステムが提供
- データアーカイブシステムと連携したDMPシステムを提供研究後のデータアーカイブ領域の自動確保
 - ResDataで作成したDMP情報に基づいてUNSW運用のストレージ「DataArchive」をリクエスト可能
 - 主に研究後の研究データ保管が円滑に行われるよう設計

UNSW : RDMoT



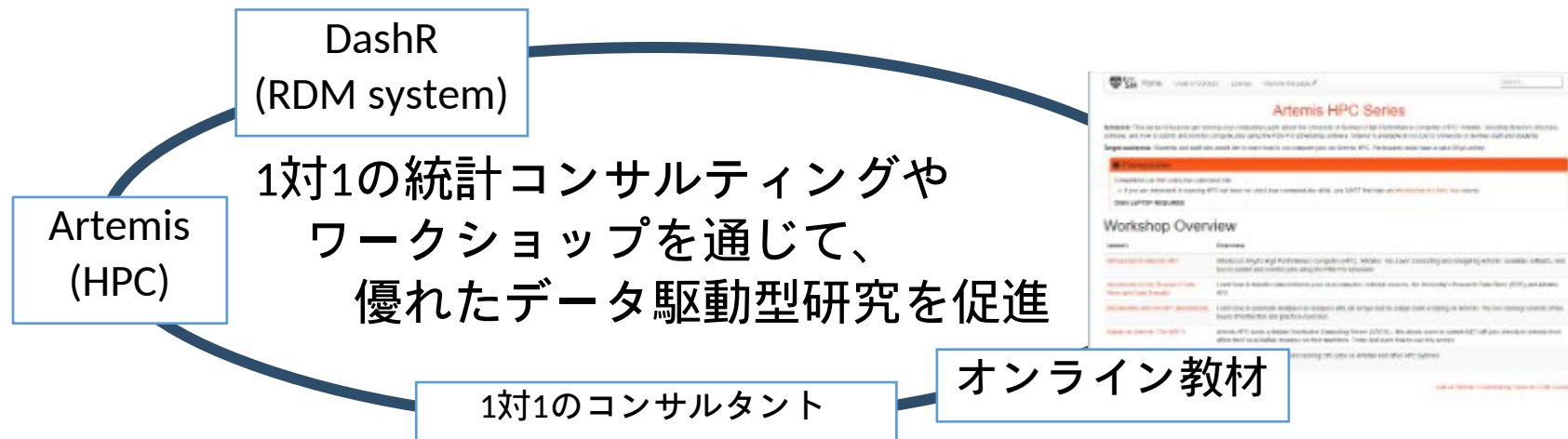
- ・ RDMオンライン教材を提供
- ・ LMSを導入して独自にオンライン教材を作成
- ・ RDMの一般知識と併せて学内ツールのインストラクションも 実施
- ・ 学習履歴の分析：離脱率や回答の流れを集計・可視化
- ・ Hacky Hour：RDMを含めた研究
ツールのレクチャーとワークショップ
- ・ ResBaz：国際的なHacky Hourとしてのカンファレンス（研究者、大学院生のマッチングメイキングの役割もある）

USYD : DashR



- DashR : DMP作成とストレージ自動確保
- UQ RDMと類似したアーキテクチャ
- HPCと連携を自動化できる
- 対応ストレージサービスはまだ少ない（今後拡大予定）
- IR機能も実装
 - どの学部の間でコラボレーションがあるか？
 - どの学部がどの程度ストレージを使用しているか？
 - 実際に経営層が使用（ストレージ予算の説明など）

USYD : Sydney Informatics Hub



- データサイエンスのトレーニングワークショップ
 - Artemis(USYDのHPC)を利用した研究のインストラクション
 - 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
 - DashRを使用した研究データ管理のインストラクション
- Hacky Hour
 - 研究者やHDRが気軽に情報交換や相談ができるセッションを開催しデータ駆動型研究の実践コミュニティを構築

まとめ方としての要点

- データセントリック時代に向けて
 - 研究力を強化するために組織として対応
 - 重要性の認識 → 対応
- 対応策
 - 組織改革
 - ITインフラの整備
 - トレーニングコースの充実
 - 学内の連携状況や研究状況などを分析
- 位置づけ
 - データセントリックに必要な解析ツールやスキルだけではなく、研究データ管理やオープンサイエンスのことも含めて対策
 - 研究力強化に繋がっていることの評価まで実施