

# RDA参加報告

報告会「研究データ共有によるイノベーションの創出  
～第8回RDA総会等の国際議論を踏まえて～」

平成28年10月3日 国立国会図書館

物質・材料研究機構 技術開発・共用部門

科学情報プラットフォーム エンジニア

田辺 浩介 TANABE.Kosuke@nims.go.jp

# 自己紹介

- 物質・材料研究機構において研究成果や情報の発信と利活用に関する業務に従事
  - 図書・電子ジャーナル
  - 機関リポジトリ
  - 研究者総覧
  - 上記に関する情報システムの管理・開発・運用
- RDAは今回初めて参加

The screenshot displays the NIMS Digital Library website. At the top, it features the NIMS logo and the text 'Library of Materials Science' and '材料科学専門図書館'. Navigation links include 'English', 'ログイン', and 'A A A'. The main content area is divided into several sections:

- 資料の検索 (Search):** A search box with a search button and a list of search criteria: '例) セラミックス 金属, セラミックス OR 金属 (詳細検索)'. Below it are links for 'ジャーナル・論文だけを探す' and '購読している電子書籍を探す'.
- お知らせ (Notice):** A tweet from @nims\_library stating that NIMS's network services will be interrupted from 3/18 to 3/20.
- 新着図書 (New Arrivals):** A list of books with their titles, publishers, and UDC numbers. Examples include '機能材料 2016年3月号: 特集イオン液体をデザインする: 多機能化への挑戦 36巻 3号' and 'Essentials of writing biomedical research papers'.
- お知らせ (Notice):** A tweet from @nims\_library regarding a power outage on 11/22 (Friday) from 5 PM to 11/24 (Tuesday) 9 AM.
- Right Sidebar:** A vertical stack of database logos including NIMS Papers Database, SAMURAI, Science and Technology of Advanced Materials, MatNavi, WEB OF SCIENCE, Journal Citation Reports, Essential Science Indicators, SciVerse Scopus, arXiv.org, Landolt-Börnstein, SciFinder, Alloy Phase Diagrams, ICSD, and WIPO.

# 参加したRDAのセッション

- Joint meeting of IG RDA/CODATA Materials Data, Infrastructure & Interoperability, WG International Materials Resource Registries (9 月 15 日)
- IG Education and Training on handling of research data(9 月 16 日)
- WG Data Security and Trust BoF(9 月 17 日)



# Joint meeting of IG RDA/CODATA Materials Data, Infrastructure & Interoperability, WG International Materials Resource Registries

- 材料科学のデータ共有に関するセッション
- Materials Resource Registryの紹介
- WGの今後の活動内容の説明

# Materials Resource Registry

- NIST（アメリカ国立標準技術研究所）が構築している材料科学のデータリソースのレジストリ
- 収録対象は材料科学に関するデータリソース（データセット、データベースサービス、ソフトウェアなど）
- ユーザ登録を行って自分でリソースを追加することも可能

The screenshot shows the Materials Resource Registry website. At the top, there is a navigation bar with links for Home, Services, Login, Help, and Contact. The main header features the NIST logo and the text "An NDS Pilot Project in collaboration with NIST". Below this, the title "Materials Resource Registry" is displayed with a "Beta" badge. A search bar is present with the placeholder text "Enter keywords, or leave blank to retrieve all records". Below the search bar, it indicates "76 results" and provides a grid of icons for different resource types: All Resources, Organizations, Data Collections, Datasets, Services, Informational Sites, and Software. A "Brief Results View" dropdown is visible. The main content area displays a list of resources, including "Materials Data Curation System: NIST Collaboration Sandbox" and "Materials Virtual Lab". Each entry shows the publisher and subject information, along with buttons for "Resource Details" and "Go To".



## 検索結果の例

- NIMSの材料データベース「MatNavi」  
(<http://mits.nims.go.jp>)  
がすでに登録されている
- 組織情報・データベース・リポジトリについて、それぞれ10件強の登録あり
- データセットの登録はまだ1件（2016年9月末時点）

The screenshot shows a web application interface with a navigation bar at the top containing tabs for 'All Resources', 'Organizations', 'Data Collections', 'Datasets', 'Services', and 'Informational Sites'. Below the navigation bar, there is a 'Resource Details' window. On the left side of the window, there is a 'Brief Results View' section and a 'Resource Type' list with radio buttons. The 'Resource Type' list includes 'All Resources' (selected), 'Organization', 'Data Collection', 'Repository', 'Project Archive', 'Database', 'Dataset', 'Service', 'Informational Site', and 'Software'. The 'Resource Details' window displays the following information:

Resource Details	
NIMS Materials Database	
localid	TM88H1J64HU8LZJJOI3
status	active
title	NIMS Materials Database
shortName	MatNavi
version	4.0
logoURL	<a href="http://www.nims.go.jp/research/hdfqf1000005np2-img/bnr_matnavi01.jpg">http://www.nims.go.jp/research/hdfqf1000005np2-img/bnr_matnavi01.jpg</a>

At the bottom right of the 'Resource Details' window, there is an 'Ok' button. At the bottom left of the main application area, there is a 'Clear Refinements' link.

# レジストリ構築の目的

- データドリブンな研究でのデータ共有の最初の段階としてレジストリを構築
- 各研究組織で独自にレジストリを立ち上げることを想定
  - 検索サービスの機能改善と単一障害点回避が目的
- 各レジストリ間は OAI-PMH でデータを同期
  - レジストリの実装は、NIST のものを用いても独自に用意してもよい
  - NISTの実装はGitHub（ソースコード公開・共有サービス）で公開  
<https://github.com/usnistgov/MaterialsResourceRegistry>

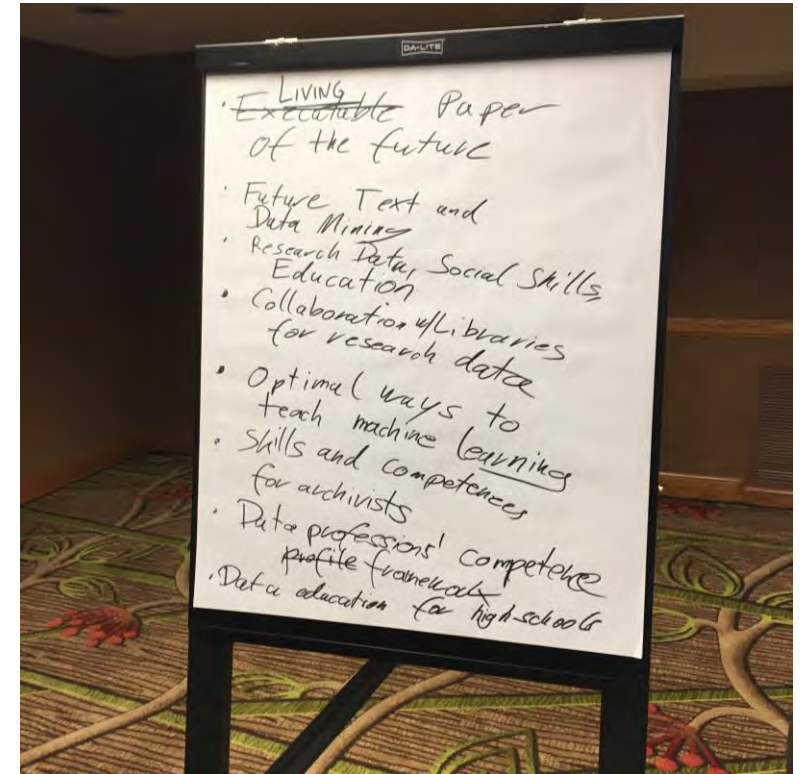
# 今後の活動内容

- 今後12ヶ月～18ヶ月をかけて、レジストリの運用により、データ共有に関する問題点を洗い出す予定
- 同時にデータ交換のためのメタデータスキーマも設計
  - Dublin Coreの拡張と材料科学特有の部分によって構成
- コミュニティ参加者には以下の協力を依頼
  - レジストリに自組織のリソースを追加する
  - スキーマの開発への協力（開発中のスキーマも[GitHubで公開中](#)）
  - 自組織でのレジストリの構築・開発・運用、OAI-PMHのサポート
  - OAI-PMH以外のデータ交換方法の提案



# IG Education and Training on handling of research data

- 研究データ管理の教育に関するワーキンググループ
  - 研究データ管理に求められる技能を特定し、教育プログラムや関連リソースを作成することを目指す
- 「テキストマイニングの教育」「機械学習の教育」「ソーシャルコミュニケーション」「図書館との協力」など、テーマごとに参加者で分かれて議論
  - 田辺は「機械学習の教育」のグループに参加



# 議論の内容

- 以下のような議論。コーディネーターは[RDA/CODATA Summer Schools](#)の講師
  - 機械学習の入門講義を2日半で実施してみたが、相当に大変。受講者のスキルの個人差も、専門の違いも大きい
  - ツールの使い方を教えるのなら難しくないが、「こういう場合の分析にはこういうアルゴリズムが適している」というのをどう教えるのか。分野や専門によってさまざまな場合があり、適切なアルゴリズムを選択できるようになるのは容易ではない
  - 使用するデータセットが受講者の専門に身近なものでないと、機械学習がどう役に立つのかわかりにくい。データセットの蓄積が必要ではないか
  - 講義を有効なものにするためには、なぜあなたの研究に機械学習のスキルが求められるのか、ということ講師と受講者が理解する必要がある

# WG Data Security and Trust BoF

- 研究データの安全な交換・共有に求められる技術的なセキュリティ要件を策定するためのワーキンググループ
- アンケート調査やワークショップの報告を受けて議論
  - 組織・職員ごとにセキュリティポリシーが異なる状態で、どのようにデータの安全性を確保するか。他のWGとの情報共有が必要
  - エンドユーザが研究活動の中でデータの検証を行えるようになっていくことが重要
  - セキュリティ情報と研究者識別子（ORCIDなど）の関係

# 番外: Sustainable Business Models for Data Repositories

- RDA ではなく [SciDataCon](#) での発表
- 世界各国のデータリポジトリのビジネスモデル（収益構造）をライトニングトーク形式で紹介
  - リポジトリの扱う分野やサービス対象ユーザーによって、課金方法は大きく異なる
  - [tDAR](#) の「メタデータ付与の代行など、Full-managed なサービスの提供で利用が大幅に増加」という発表が興味深かった

# 収入源によるビジネスモデルの比較

主な収入源	長所	短所	例
国からの交付金	比較的安定した収入が見込める	収入増が困難 プロジェクトによる補助金の場合、長期運用に不安がある	Digital Repository of Ireland など
大学への課金	比較的安定した収入が見込める	収入増が困難	The Arabidopsis Information Resource など
企業への課金	大規模な収入が見込める	企業の業績や経済状況によって収入が不安定になる	The Cambridge Structural Database など
個人への課金	幅広い利用者を対象にできる	多数の利用者を集めなければ運営が困難	?

これらの収入源をどのように組み合わせて運用するか？

# 情報統合型物質・ 材料開発イニシア ティブ (MI2I)

- 材料科学研究のための  
データプラットフォーム  
の構築を行う
- 企業・大学が経費を持ち  
寄るコンソーシアム形式

ホーム  
HOME

MI<sup>2</sup>I について  
ABOUT

各グループの説明  
GROUP INFORMATION

イベント  
EVENT

ニュース  
NEWS



## What's New

2016.09.08

第三回 MI<sup>2</sup>Iコンソーシアムイベント **NEW**

2016.07.15

第二回 MI<sup>2</sup>Iコンソーシアムイベント **NEW**

2016.07.13

コンソーシアムキックオフが開催されました **NEW**

## 各グループの説明

情報統合型物質・材料研究材料研究拠点  
蓄電池材料グループ

情報統合型物質・材料研究材料研究拠点  
磁石・スピントロニクス材料グル  
ープ

情報統合型物質・材料研究材料研究拠点  
伝熱制御・熱電材料グループ

情報統合型物質・材料研究材料研究拠点



# 参加した感想

- 「自分の組織でやったことを持ち寄って合意を形成する」という進め方が面白い
  - （スキーマでもシステムでも）合意を形成するには時間がかかるが、合意が決まってから自分の組織で取り組み始めるのでは取り残される
- 「研究支援職（図書館員など）の役割は、自組織の構成員にデータ管理の重要性を伝えること」という意識を強く感じる
  - 「自分がどう勉強するか」という話が先に立ってしまいがちだが...
- システムにしてもポリシーにしても、単一の正解はなさそうなので、怖がらずに始めたい

# エンジニアの視点から

- 共有されているデータを使用する際、データへの「信頼」をどう担保するのか？
  - プログラムのソースコードを動作する環境で提供する例もある  
ex. IPOL Journal (<http://www.ipol.im>)
- 究極的には論文やデータ、あるいは著者に対する電子証明書が必要？
  - Webサイト（SSL証明書）やソフトウェアの配布のような...
  - 現在の著者ID（ORCID, ResearcherIDなど）では、本人証明にはならない