

第3回 SPARC Japan セミナー2016

「科学的知識創成の新たな標準基盤へ向けて：オープンサイエンス再考」

RDM トレーニングツールの紹介

尾城 孝一

(東京大学附属図書館)

講演要旨



世界各国で公的研究資金による研究成果について、論文だけではなくエビデンスとなる研究データの保存や共有を求めるポリシーを制定する例が相次いでいる。適切に研究データを管理するには、研究者自身に加えて研究支援者が研究データ管理に係る知識やスキルを習得する必要がある。本発表では、機関リポジトリ推進委員会の研究データタスクフォースで開発を進めている、研究データ管理に関する研究支援者向け学習教材について紹介する。



尾城 孝一

東京大学附属図書館事務部長。1983年1月、名古屋大学附属図書館に採用され図書館職員としてのキャリアを開始。その後、東京工業大学附属図書館、国立国会図書館、千葉大学附属図書館、国立情報学研究所、東京大学附属図書館、大学図書館コンソーシアム連合事務局を歴任。2015年4月より、現職。

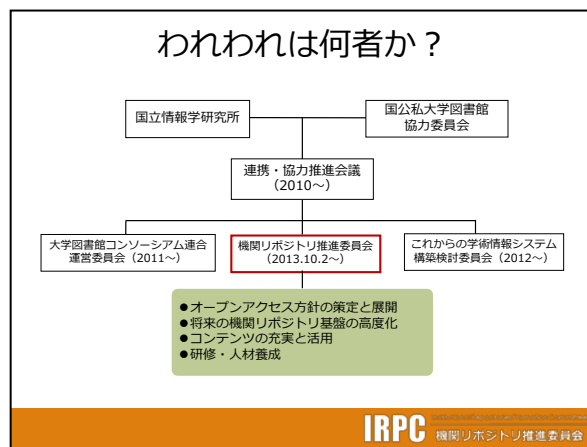
今日は、機関リポジトリ推進委員会の研究データタスクフォースが開発した、RDM (Research Data Management) トレーニングツールについて紹介させていただきます。

おり、われわれは其中で研究データタスクフォースとして活動しています。これから紹介する RDM トレーニングツールの開発や、ケーススタディを通じて研究データ管理のノウハウを蓄積していくような活動をしています。

機関リポジトリ推進委員会の取り組み

国立情報学研究所 (NII) と大学図書館の連携・協力の枠組みの中に、機関リポジトリ推進委員会という委員会が置かれており、そこで機関リポジトリ、オープンアクセス、オープンサイエンスに関するさまざまな活動をしています (図 1)。

機関リポジトリ推進委員会には、基幹的な業務のサポートをするための三つの常置の作業部会があります (図 2)。その他、先端的なプロジェクトに取り組むためのアドホックなタスクフォースが六つ設置されて



(図 1)

図3がそのタスクフォースのメンバーです。私の他に、NIIの山地先生、大学や研究所の図書館員の皆さん、URAの皆さんにも入ってもらい、一緒に活動しています。数は少ないのですが、これは今考えられる最強の精鋭部隊だと自負しています。

さて、オープンサイエンスなのですが(図4)、研究データに対して研究者の皆さんはどういう意識を持っているのでしょうか。DataONEというプロジェクトに関連する調査があり、図5はその最初の方の調査結果なのですが、1,329名の研究者のうち、78%が、少なくともデータの一部はアクセス制限のないデータリポジトリに登録したいと考えている、しかし、自分自身のデータに他の研究者が容易にアクセスできると回答した回答者は36.2%にとどまっています。データを電子的に利用できるようにできない理由としては、やはり時間がない、お金がない、データの登録先

がないといったことが挙げられています。

この調査から、データを公開したい、あるいは研究資金配分機関等のポリシーによって公開しなければならない、けれども実現できないでいる研究者が結構たくさんいるということがうかがえるかと思えます。

こういう研究者を支援し、オープンサイエンスを進めていくためには三つの基盤が必要だと考えています。一つ目は研究データ基盤(共通システム)、二つ目は研究者を支援する人の基盤(支援スタッフの確保・育成)、三つ目は組織の基盤(コミュニティ)です。今日はこのうち二つ目の支援スタッフの確保・育成につながるトレーニングツールの紹介をします。

RDM トレーニングツールとは

文部科学省が2016年2月に公開した審議のまとめの中でも、研究データのオープン化に関して、大学等

機関リポジトリ推進委員会

常置の作業部会 (基幹業務のサポート)

JMIR Cloud 運用作業部会	広報作業部会	研究作業部会
<ul style="list-style-type: none"> • ICの安定運用と機能改善 • ICへの移行サポート • ICコミュニティ支援 • SCPIの移行準備 	<ul style="list-style-type: none"> • イベント、説明会等の開催 • Web, facebookの編集・管理 • 国際協力のための海外派遣 	<ul style="list-style-type: none"> • 担当者研修

アドホックなタスクフォース (先端的なプロジェクト)

研究データTF	論文OA TF	メタデータTF	指標・評価・メトリクスTF	ORCID TF	CoAR Asia TF
<ul style="list-style-type: none"> • RDM上レベリングツールの開発 • クラウドベースによる研究データ管理ソフトウェアの開発 	<ul style="list-style-type: none"> • オープンアクセスの針決定支援 • ポリシー決定後の実施支援 • OA評価・トラッキング 	<ul style="list-style-type: none"> • metadataに関する新たなメタデータスキーマの開発 • 研究リポジトリへの適用・実装支援 	<ul style="list-style-type: none"> • IRISコンテンツランキングの公開 • アクセス統計提供システム • アクセスログ解析 	<ul style="list-style-type: none"> • ORCIDのコンソーシアムライセンスを包含して提供することを想定し、情報収集と検討を行う。 	<ul style="list-style-type: none"> • AsiaでOAのコミュニティを形成

IRPC 機関リポジトリ推進委員会

(図2)

オープンサイエンスとは

オープンサイエンスとは、オープンアクセスと研究データのオープン化(オープンデータ)を含む概念である。オープンアクセスが進むことにより、学術・産業界、市民等あらゆるユーザーが研究成果を広く利用可能となり、その結果、研究者の所属機関、専門分野、国境を越えた新たな協働による知の創出を加速し、新たな価値を生み出していくことが可能となる。

第5期科学技術基本計画、本文p.32
www.8.cao.go.jp/cstp/kihonkeikaku/Shonbun.pdf

IRPC 機関リポジトリ推進委員会

(図4)

タスクフォースメンバー

主査) 尾城 孝一 東京大学附属図書館
 副主査) 山地 一禎 国立情報学研究所学術ネットワーク研究開発センター、学術リポジトリ推進室

前田 翔太 北海道大学附属図書館
 三角 太郎 千葉大学附属図書館
 天野 絵里子 京都大学学術研究支援室
 大園 隼彦 岡山大学附属図書館
 西園 由依 鹿児島大学学術情報部
 南山 泰之 情報・システム研究機構 国立極地研究所情報図書室

IRPC 機関リポジトリ推進委員会

(図3)

研究者の意識は?

- 78%の回答者が、少なくともデータの一部はアクセス制限のないデータリポジトリに登録したいと考えている。
- しかし、自身のデータに他の研究者が容易にアクセスできるとした回答者は36.2%に留まっている。
- データを電子的に利用できるようにできない理由は、
 - 時間がない (53.6%)
 - お金がない (39.6%)
 - データの登録先がない (23.5%)

Tenopir, Carol. et al. Data sharing by scientists: practices and perceptions. Plos ONE. 2011, 6(6): E21101.
<http://dx.doi.org/10.1371/journal.pone.0021101>

→ データを公開したい、あるいは公開しなければならないが、実現できずにいる研究者が多数存在する

IRPC 機関リポジトリ推進委員会

(図5)

に期待されている役割が幾つか挙げられているのですが、その一つとして、技術職員、URA 及び大学図書館職員等を中心としたデータ管理体制を構築し、研究者への支援に資するということが書かれています (図 6)。

しかし、研究データ管理の支援をやれと言われても、一体何をやっていいのかよく分かりません。これは海外でも事情はどうも同じようで、どういうことをやればいいのか、それが学べるオンラインのツールがたくさんできています (図 7)。その日本版の教材をつくらうというのが、われわれが RDM トレーニングツールの開発に取り掛かったきっかけです。

ツールの目的は、学習者が RDM に関する基礎的な知識を得ること、研究者支援サービスの足掛かりを得ることです。

教材の形式は、eラーニング教材にして、MOOC で

提供できるようにします (図 8)。講義、確認テスト、ワークなども適宜挟みこんでいきます。受講者は最初は図書館員を想定していたのですが、せっかくならもっと幅広く、IT 部門の人や、URA といった研究支援部門の人たち、それからオープンサイエンスや研究データに関心を持っている人たちに使ってもらえるようなツールにできればということで、教材の開発に取り掛かって、約 1 年かかってほぼ完成に近づきつつあります。

全体の教材の構成は図 9 のようになっており、7 章から成っています。データの生成、加工、分析、保存、公開、再利用と、データのライフサイクル全般についてひとつひとつ学べるようになっていきます。

北本先生から、2017 年のスキルと 2027 年のスキルというお話がありましたが、取りあえずこのツールは 2017 年のスキルについて学べるようになっていきます。

研究者支援への期待

『学術情報のオープン化の推進について (審議まとめ)』
平成28年2月 科学技術・学術審議会 学術分科会 学術情報委員会
http://www.mext.go.jp/b_menu/shing/gijyutu/gijyutu4/036/houkoku/1368803.htm

(大学等に期待される役割のひとつ)

技術職員、URA及び大学図書館職員等を中心としたデータ管理体制を構築し、研究者への支援に資する

IRPC 機関リポジトリ推進委員会

(図 6)

教材の概要

- eラーニング教材 (MOOCにて提供予定)
 - 講義、確認テスト、ワーク
- 受講者は幅広く想定
 - 大学や研究機関等において研究者の支援に携わる方 (図書館・IT部門・研究支援部門の職員 (URA含む)、研究室で研究補助業務に携わる方等)
 - オープンサイエンスや研究データ管理に関心のある方

IRPC 機関リポジトリ推進委員会

(図 8)

研究データ管理の教材

(海外の教材)

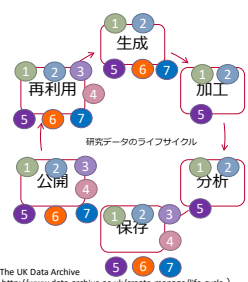


→ 日本版の教材を作ろう!

IRPC 機関リポジトリ推進委員会

(図 7)

データのライフサイクルと教材の構成



- ① 1章 導入
- ② 2章 データ管理計画
- ③ 3章 保存と共有
- ④ 4章 組織化・メタデータ
- ⑤ 5章 法倫理的問題
- ⑥ 6章 ポリシー
- ⑦ 7章 サポートサービスの検討

* The UK Data Archive (<http://www.data-archive.ac.uk/create-manage/life-cycle>)
 * RDM Support - basic training course for information specialists (<http://dx.doi.org/10.6084/m9.figshare.1285313>) より

IRPC 機関リポジトリ推進委員会

(図 9)

若い人にとっては、これは人生戦略上の損失につながってしまうようなツールかもしれないのですが、まずは基礎的なところから学んでいくのも大事ではないかということです。

図 10・図 11 に、各章のシラバスを簡単に書いてあります。導入、データ管理計画、保存と共有、組織化・メタデータ、法倫理的問題、ポリシー、サポートサービスの検討、このような内容になっています。

例えば第 1 章では、研究データ管理とは何かという説明があり、それに関するワークなども挟みこまれています (図 12・図 13)。

また、JST の小賀坂さんからデータ管理計画 (DMP) のお話がありましたが、DMP のつくり方なども学べるようになっています (図 14・図 15)。

今後の展開

この教材は、まず 3 月末までに完成させて、MOOC 化も併せて完了させて公開したいと考えています (図 16)。その後、4 月から、JPCOAR (オープンアクセスリポジトリ推進協会) というコミュニティが本格的に活動を始めます。今の機関リポジトリ推進委員会の活動も、この協会の活動に引き継がれていきます。われわれ研究データタスクフォースもこの新しい協会の中で、教材をより広く活用してもらって、学んだ人からフィードバックを得て改定して、より良いサービスにしていきたいと考えています。

2027 年のスキルについて一緒に考えて、こういう活動を共にやっていきたいという方がいらっしゃいましたら、ぜひ声を掛けていただければと思います。

シラバス

第1章 導入
研究データ管理の重要性が増している背景や、研究データ、研究データ管理の定義について

第2章 データ管理計画
効果的なデータ管理に欠かせないデータ管理計画に関し、作成義務化の動向や、その構成要素について

第3章 保存と共有
研究データの、研究期間中の保管や長期保存に関する留意点について。研究データの共有に関し、その意義や検討すべき点、共有方法について。

第4章 組織化・メタデータ
研究データを長期的に管理・活用するために欠かせない、一定のルールに則ったデータの組織化や、データについて説明する文書やメタデータの作成について。

IRPC 機関リポジトリ推進委員会

(図 10)

サンプル (未定稿)

1.3.1 研究データ管理 (Research data management) とは

ここでいう「研究データ管理」という言葉は、ある研究プロジェクトにおいて使用された、あるいは生成された情報をどのように組織化し、保存し、管理していくのかといったことにも至る。

研究データ管理には次のようなことが含まれる: > 第2章も参照

- どのようデータを取り扱っていくかの計画 (データ管理計画/DMP) - 多くの助成機関においてデータ管理計画が申請書類の一部として求められている
- 研究プロジェクトが行われている間の、日々の情報の取り扱い方
- 長期的にデータを何に使うのか (研究プロジェクトが終わった後はそのデータを何に使用するか) <http://researchdata.ox.ac.uk/home/introduction-to-rdm/> より

IRPC 機関リポジトリ推進委員会

(図 12)

シラバス (続)

第5章 法倫理的問題
研究データをめぐる著作権や、再利用を促進するためのライセンスの仕組みについて。センシティブデータを取り扱う上での留意点や研究倫理について。

第6章 ポリシー
国や助成団体、機関、出版者等による、研究データの保存や共有を求めるポリシー策定の動向や要件について

第7章 サポートサービスの検討
前章までの内容を踏まえ、学習者が自機関での研究データ管理サービスを構築していくためのステップについて

IRPC 機関リポジトリ推進委員会

(図 11)

サンプル (未定稿)

1.4.1 ワーク (基礎)

次の文章の空欄に相応しい言葉を、選択肢a~fの中から選択してください。

▶ オープンサイエンスとは論文へのオープンアクセスと「 」のオープン化を含む概念であり、我が国では内閣府の出した第五期科学技術基本計画でその取り組みの姿勢について言及された。

▶ 研究データはオープンサイエンスや「 」の観点からも適切な保存や公開が求められている。

▶ 研究資金の助成機関では「 」の提出が義務付けられていることもある。

▶ 研究データ管理のことを英語で「 」という。そのメリットとして、「 」や「 」などが挙げられる。

選択肢:
a データ消失リスクの減少 b Research Data Management (RDM)
c 研究データ d 研究効率の上昇 e Data Management Plan (DMP)
f 不正対策

IRPC 機関リポジトリ推進委員会

(図 13)

サンプル（未定稿）

2.1.1 データ管理計画（DMP）とは

- ▶ 「研究プロジェクト等における研究データの取り扱いを定めるものであり、具体的にはデータの種類、フォーマット、アクセス及び共有のための方針、研究成果の保管に関する計画などについて記載されるもの」（文部科学省科学技術・学術審議会学術分科会学術情報委員会（2016）「学術情報のオープン化の推進について（審議のまとめ）」用語解説より）
- ▶ 研究の実施段階から終了後に至るまでの期間において、研究データがどのように生成、管理、共有、保存される予定か、を文書化する。
- ▶ この計画に従った管理を行い、研究データが利活用可能な状態で適切に管理されるようにする。
- ▶ 研究データの適切な保管・管理は、研究データの公開を進めるための前提である。

IRPC 機関リポジトリ推進委員会

(図 14)

サンプル（未定稿）

2.4.1 データ管理計画理解度チェック

1. 次の文章の空欄に相応しい言葉を、選択肢a～fの中から選択してください。

データ管理計画は以下に挙げるようなことで役立つ：

1. 自分の研究データや記録が、[]であること、完全であること、真正性があること、信頼できること、これらのことを保証する（ラベリングが意味のあるもので完全性があるということを保証するのに役立つ）..
2. 長期的に見て、[]と資源の節約になる。
3. データの[]を高め、そのことにより、データ損失の[]を最小化する。
4. 研究の公正性と、他者による[]を保証する。
5. 他者に自分のデータを利用させることにより、手間の[]を防ぐ。

選択肢：
a. セキュリティ b. 再現性 c. 正確 d. 重複 e. 時間 f. リスク

IRPC 機関リポジトリ推進委員会

(図 15)

今後の展開

- 教材の公開（3月末予定）

↓

J P C O A R
オープンアクセスリポジトリ推進協会

2017.4 本格始動

- 教材の活用促進
- 持続可能な教材開発体制の確立
- 教材の段階的拡充

IRPC 機関リポジトリ推進委員会

(図 16)