

研究データ管理の理解促進と支援 サービス構築に向けた教材開発の 取り組み

尾城孝一¹⁾, 西園由依²⁾, 天野絵里子³⁾, 大園隼彦⁴⁾, 前田翔太⁵⁾, 三角太郎⁶⁾,
南山泰之⁷⁾, 山地一禎⁸⁾

^{1) 8)} 国立情報学研究所オープンサイエンス基盤研究センター, ²⁾ 鹿児島大学学術情報部,

³⁾ 京都大学学術研究支援室, ⁴⁾ 岡山大学附属図書館, ⁵⁾ 北海道大学附属図書館,

⁶⁾ 筑波大学附属図書館, ⁷⁾ 国立極地研究所

Who we are

オープンアクセスリポジトリ推進協会 (Japan Consortium for Open Access Repository: JPCOAR)

- オープンアクセスの発展に向けた課題の解決に共同で取り組む会員制のコンソーシアム
- 2016年7月に発足
- 全国の500を超える大学図書館等が参加
- 活動（3つの作業部会と4つのタスクフォース）
 - オープンアクセス基盤の安定的運用
 - オープンアクセスの先端的機能の開発
 - オープンアクセスの発展に寄与する国際連携

研究データタスクフォース

- RDMトレーニングツール（研究データ管理のオンライン教材）の開発と利用促進
- ケーススタディによる研究データ管理ノウハウの蓄積

2016年度メンバー

尾城孝一	東京大学附属図書館（主査）
山地一禎	国立情報学研究所（副主査）
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	国立極地研究所情報図書室

2017年度メンバー

山地一禎	国立情報学研究所（主査）
尾城孝一	国立情報学研究所
常川真央	千葉大学附属図書館
松本侑子	東京大学附属図書館
青木学総	京都大学情報環境機構
天野絵里子	京都大学学術研究支援室
西園由依	鹿児島大学学術情報部
南山泰之	国立極地研究所情報図書室

研究データ管理とは

研究データ管理（Research Data Management: RDM）とは、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データの組織化、構造化、保存、共有、公開、再利用に関する一連の作業を指す言葉

研究データ管理の背景

1. オープンサイエンスの推進

- 論文のみならず、論文のエビデンスとしての研究データも含めてオープン化を進め、イノベーションの創出を促す

2. 研究公正（不正防止）

- 論文の根拠となったデータは、論文発表後10年間保存し、必要に応じて開示

研究者と研究データ管理

- 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

エジンバラ大学のRDMサービス

4つのサービスで研究者をサポート



BEFORE

create a data
management
plan

DURING

working with
data

AFTER

share and
archive your
data

TRAINING
&
SUPPORT

(研究前)

データ管理計画
作成ツール

(研究中)

ストレージ
バックアップ
ラボノート管理

(研究後)

共有・公開
長期保存

(教育・支援)

オンライン教材
リテラシー教育
トレーニングコース
ヘルプデスク

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

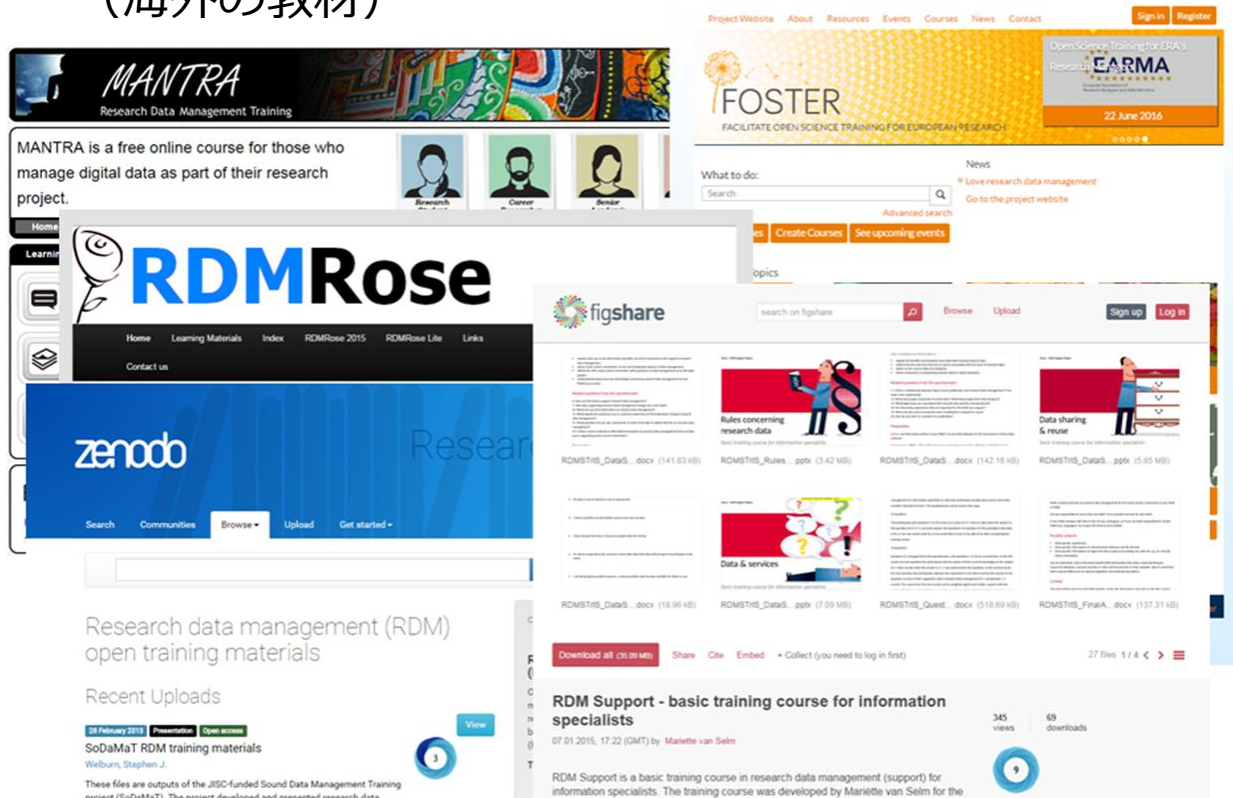
http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/036/houkoku/1368803.htm

（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

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！

教材の概要

- 目的

- 学習者が研究データ管理に関する基礎的な知識を習得する
- 研究データ管理サービス構築の足掛かりを得る

- 内容

- 全7章から構成されるeラーニング教材
- 各章は、スライドと解説で提供し、随所に確認テストを折り込む

- 受講対象者

- 図書館、IT部門、研究支援部門の職員（URA）等
- オープンサイエンスや研究データ管理に関心のある教職員

教材の公開

ログイン

J P C O A R
オープンアクセスリポジトリ推進協会

概要 About 組織 Organization 参加機関 Members 参加申込 Join us

JAIRO Cloud とは? What is "JAIRO Cloud"? 規程類 Rules 会議資料 Minutes 事業計画 事業報告 Project 各種資料 Publication

WEKO

トップ ランキング

品 検索

詳細検索 全文検索 キーワード検索

Language 日本語

インデックス ツリー

- 入会申請書
- 設立趣意書
- 規程類
- 設立総会
- リーフレット
- 発表資料等
- 事業計画・事業報告
- 作業部会
- タスクフォース
- 論文OA
- 研究データ

インデックス

事業計画・事業報告 タスクフォース 研究データ

Permalink: <http://id.nii.ac.jp/1458/00000023/>

RDMトレーニングツール

利用統計は来月からご利用いただけます

File / Name	License
RDMトレーニングツール(スクリプト入)_1章	
RDMトレーニングツール(スクリプト入)_1章 (166.56KB) [6 downloads]	Creative Commons : 表示
RDMトレーニングツール(スクリプト入)_2章	
RDMトレーニングツール(スクリプト入)_2章 (803.53KB) [2 downloads]	Creative Commons : 表示
RDMトレーニングツール(スクリプト入)_3章	
RDMトレーニングツール(スクリプト入)_3章 (118.66KB) [1 downloads]	Creative Commons : 表示
RDMトレーニングツール(スクリプト入)_4章	
RDMトレーニングツール(スクリプト入)_4章 (454.65KB) [1 downloads]	Creative Commons : 表示
RDMトレーニングツール(スクリプト入)_5章	
RDMトレーニングツール(スクリプト入)_5章 (578.65KB) [1 downloads]	Creative Commons : 表示
RDMトレーニングツール(スクリプト入)_6章	
RDMトレーニングツール(スクリプト入)_6章 (100.51KB) [1 downloads]	Creative Commons : 表示
RDMトレーニングツール(スクリプト入)_7章	
RDMトレーニングツール(スクリプト入)_7章 (378.61KB) [1 downloads]	Creative Commons : 表示
RDMトレーニングツールスクリプト	
RDMトレーニングツールスクリプト (61.26KB) [3 downloads]	Creative Commons : 表示

OAI-PMH BibTeX OWLSRDC WEKO REPOSITORY PRINT

アイテムタイプ その他 / others

言語 日本語

戻る

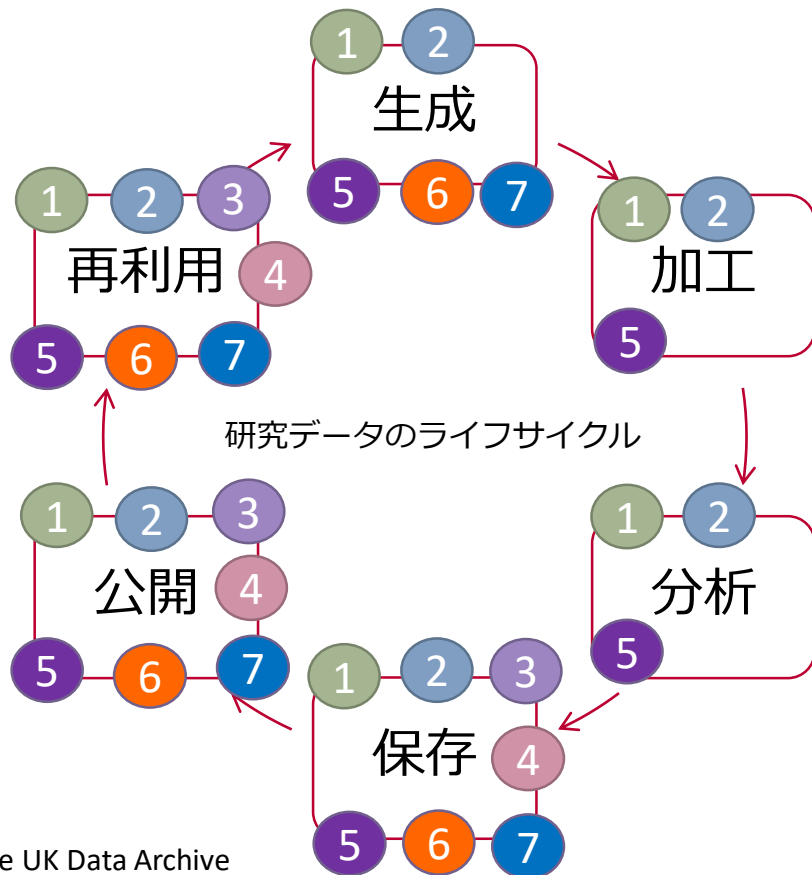
Powered by WEKO

RDMトレーニングツール

2017年6月6日にスライドとスクリプト（説明の文章）をCC BYで公開

<http://id.nii.ac.jp/1458/00000023/>

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 サービスの設計

- 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

第1章 導入

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

1.2.1 研究データとは

- ▶ 研究データとは研究の収集・生成される情報されたり、結論を導く結果が正しいかどうか。
- ▶ 電子的なデータのみ研究の情報も研究データ。
- ▶ 研究データのとりまじり、記録されるものも多岐にわたる。

1.2.3 研究データの具体例

- ▶ 観測データ (Observational)
 - ▶ 自然現象等を観測し取得されたデータ。
例：天体観測、気象観測、社会調査
- ▶ 実験データ (Experimental)
 - ▶ 実験によって得られたデータ。高価になる。
 - ▶ 例：DNAシーケンス、クロマトグラム
- ▶ シミュレーション (Simulation)
 - ▶ モデルやコードを用いて生成されたデータ。
 - ▶ モデルやコード、パラメータセットも重要。
 - ▶ 例：気象モデル、経済モデル
- ▶ 派生データ・編集されたデータ (derive)
 - ▶ 取得したデータを加工、あるいは結合する。
 - ▶ 例：データマイニング、テキストマイニング
- ▶ 参照/標準データ (Reference or canonic)
 - ▶ 研究の過程で参照されるデータ、データクレンジング。整理・共有されているもの。
 - ▶ 例：遺伝子配列データベース、結晶データ

1.3.2 研究データ管理のメリット

- ▶ 研究効率
 - ▶ 研究効率の向上
 - ▶ 長期的に時間や資源を節約
 - ▶ データセキュリティの強化とデータ消失リスクの最小化
 - ▶ データを他の研究者が再利用することにより、労力の重複を避ける
 - ▶ ある時点のデータを他の時点と比較できるデータを提供することにより、データの変化の分析を容易にする
- ▶ ポリシー
 - ▶ 資金提供元の要求を満たす
 - ▶ 産業界や商業界の慣行に従う
- ▶ 研究倫理
 - ▶ 研究の再現性の確立
 - ▶ 研究データや研究記録の正確性、完全性、真正性、信頼性の保障

J P C O A R
オープンアクセスリポジトリ推進協会

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
- 作成の義務化の動向や、その構成要素について

2.1.1 データ管理計画（DMP）とは

- 「研究
である
委員
め」
- 研究
がど
- この
切に
- 研究
の前

2.1.2 データ管理計画策定の意義

- 計画を策定することにより、適切なデータ管理を行える。適切なデータ管理を行える。適切なデータ管理を行える。
- 研究の公正性と再現性を保証する
- 自分の研究データや記録が、正しさを保証する。
- データのセキュリティを高め、
- 研究の初期段階から、データ管理とで、データ管理に関する入念
- これらは研究の質や効率性を高
- 助成機関等の意向（2.2.1参照）に不可欠なものとしてとらえらるべき。

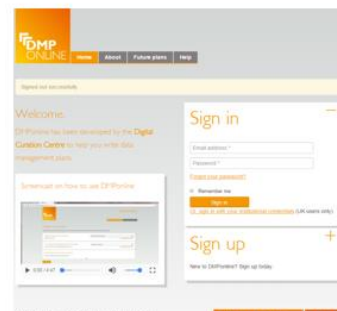
2.3.1 何をどう計画するのか

- チェックリスト（Digital Curation Centre (DCC)）
<http://www.dcc.ac.uk/resources/data-management-plans/checklist>

ツール例

- DMPTool（カリフォルニア大学）
<https://dmptool.org/>
- DMPOnline（DCC）
<https://dmponline.dcc.ac.uk/>

誰でも無料で
アカウント作成
可能



すること

J P C O A R
オープンアクセスリポジトリ推進協会

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
- 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

3.1.1 保存場所

- 研究資料(研究データ)を失うことは研究の進展に大きく影響
- ファイルの意図しない削除
- ハードウェアの故障
- モバイル機器の紛失

3.1.3 セキュリティ

- ▶ 物理的なセキュリティ対策
 - ▶ 離席する際のPCをロック
 - ▶ 部屋の施錠
 - ▶ モバイルPCを放置しない
 - ▶ USBメモリ等を紛失しない
- ▶ PC上でのセキュリティ対策
 - ▶ アンチウイルスソフトウェア
 - ▶ ソフトウェアやOSのアップデート
 - ▶ 適切な強度のパスワード
 - ▶ ファイルやハードディスク
 - ▶ 例 WindowsのBitLockerなど、機能のセキュリティ
- ▶ 個人情報・プライバシー
 - ▶ 個人情報やプライバシーを守る取り組みの方法を明確にする
 - ▶ 例 高等教育機関向けの基準(C2103 情報格付けの基

3.1.2 バックアップ

3.3.1 オープン／クローズ

- ▶ **研究データを公開することの意義**
 1. 研究成果の透明性や公正性の確保
 2. 研究資金助成機関からの要求への対応
 3. 研究成果への直接的な引用や再利用によるインパクト
 4. 新たな研究コミュニティの創出や産学連携への発展
 5. 派生的な成果や新たな解釈への発展によるイノベーションの創出
 6. 適切に管理し保存することによる自身のための再利用性の向上
 7. 若手研究者や学生に対するデータの収集や解析に関する教育的効果
 8. 市民や納税者からの研究に対する理解の向上
- ▶ **研究データを非公開とする必要性**
 1. 知的財産的な価値がある研究成果に対する戦略的な占有
 2. 財政の関係者が知的財産権を保持したり秘匿義務がある場合
 3. 個人情報やプライバシー情報などのセンシティブな情報を含む場合

単にオープン化を推進するのではなく、機関、研究者自身、研究コミュニティなどでオープン・クローズ戦略を十分に検討することが重要

J P C O A R
オープンアクセスリポジトリ推進協会

3.2.2 フォーマット

的な保存を考える場合、適切なファイルフォーマットを選択す

ファイルの拡張子(.pdf, .jpg)は、ファイルフォーマットに関する情報を提

アイルフォーマット

ソフトウェアと独立したフォーマット
標準化されているフォーマット
ているフォーマット

変換のリスク

トを変換することにより、ファイル内部に記録されたメタデータや文章構造などの情報を失う可能性大。

イナリの比較

ファイルは環境に依存せず開くことができ人間可読性高。

ファイルはよりリッチな情報を包埋していることが多いが人間可読性低。

J P C O A R
オープンアクセスリポジトリ推進協会

第4章 組織化、文書化、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
- 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

4.1.2 ファイルを組織化するための3つの判断基準

1.

4.1.3 ファイル命名規則の例

- ファイルの構成
- 特殊記号は
- ピリオド、

2.

3.

- 文脈の保存
- ファイル名
- 保存場所に
- 実験機器等
- ファイルの
- る。例)
- 一貫性のある
- 日付を用い

4.2.1 データについて文書で記録する意味

- 1. 自分自身の覚えのために
研究を実施している間はデータが記録されているが、時間の経過に従って忘れられる可能性がある。
- 2. 他の研究者の再利用のため
研究者が既存のデータを再活用する。例えば、先行研究の結果を結果を再現する、同様の研究

4.4.3 データ引用の方法

- 研究で参照したデータセットは、他の利用者がアクセスできるように適切な情報を含まなければならない。
- DataCite** は、以下の最小限の引用要素を推奨
Creator (PublicationYear): Title. Publisher. Identifier
- Identifier** は、DOI や URL 等の一意に識別するもの。
- 可能ならば **Version** と **ResourceType** も追加する。
Creator (PublicationYear): Title. Version. Publisher. ResourceType. Identifier
- UK Data Service** は、主題を表すタイトルに加えて、データがカバーする地理的情報と時間的範囲を示すことを推奨。

4.3.2 様々なメタデータの標準

- 一般的なもの : Dublin Core、DataCite
- 分野特有のもの : Data Documentation Initiative (DDI, 社会科学)
DCCのwebサイトで検索可能

メタデータの種類はどのようなタイプか？

分野特有のメタデータ
一般的なメタデータ

J P C O A R
オープンアクセスリポジトリ推進協会

J P C O A R
オープンアクセスリポジトリ推進協会

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
- 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

5.1.2 データの著作物性

著作物の定義：

「思想又は感情を創作的に
学術、美術又は音楽の範囲に
法第2条第1項第1号）。

- ▶ 単なる客観的事実やデータは
ない
- ▶ データを得るために高度の知
たとしても、保護対象になら

もっとも、
一定の考え方のもとにデータ
「創作性」が認められる場合
例：データベースの著作物

5.2.1 ライセンシング

- ▶ ライセンシング（利用許諾）とは
 - ▶ コンテンツの利用許諾の条件を明示すること。
- ▶ 趣旨
 - ▶ 利用許諾に関する曖昧さの回避
 - ▶ 使ってよい範囲を明確にしておくことで、再利用の
 - ▶ 各国で異なる著作権ルールを標準化
- ▶ 種類
 - ▶ Creative Commons License
 - ▶ <https://creativecommons.jp/licenses/>
 - ▶ Open Data Commons
 - ▶ <http://opendatacommons.org/licenses/>
 - ▶ Open/Non-Commercial Government Licence
 - ▶ <http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence.htm>
 - ▶ Public Domain

5.4.1 センシティブデータとは

▶ センシティブデータとは？

1. 個人情報
 - ▶ 種族・民族に関するもの
 - ▶ 政治的意見・宗教に関するもの
 - ▶ 労働組合での活動など
 - ▶ （肉体的・精神的な）健康状態に関するもの、性に関するもの
 - ▶ 守秘義務に係るもの（例：被弁護人に係る情報）
2. 国家安全保障に関するもの
3. 知的財産権
4. その他特に配慮を必要とするもの
例：絶滅危惧種の生息地情報

J P C O A R
オープンアクセスリポジトリ推進協会

J P C O A R
オープンアクセスリポジトリ推進協会

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

6.1.2 研究者が遵守を求められる三つのポリシー

- ▶ 1. 研究機関
 - ▶ 大学、研究所
- ▶ 2. 配分機関
 - ▶ 研究機関に対し
- ▶ 3. 学会・出

6.1.3 研究機関における研究データ管理の二つの流れ

▶ 1. オープンサイエンスの流れ

「我が国におけるオープンサイエンス推進のあり方について」内閣府 2016
「オープンイノベーションに資するオープンサイエンスのあり方に関する調査」
研究データの利活用のために積極的にオープンサイエンスを推進する

▶ 2. 研究公正の流れ

「研究活動における不正行為への対応等に関するガイドライン」文科省 2016
研究データは検証が目的、利活用を想定している
→研究データをオープンにして、多数に検証してもらうという考えがある
あるであろうが、多くの場合は、ポリシーの作り方の問題である

1と2では、オープン化に対してのスタンスが異なる
研究公正の流れのポリシー策定が先行。オープンサイエンスを推進するのであれば、2との整合性を保つ必要がある

6.1.6 想定される課題

- ▶ 各機関・各部局で論文公表の実態のモニタリングはできているのか？
- ▶ 研究データの管理はシステム化されているのか
 - ▶ データの検索は？（大規模大学では数千本の論文が公表されるが、その根拠データをどのように管理？）
 - ▶ 公表論文とデータの紐づけは？

【研究支援担当者側で検討できそうなこと】

○研究データ管理ポリシーの具体化をはかるためには、研究成果のモニタリングおよびデータベース化が必要になってくるであろうし、そこでオープンサイエンスとの課題の共有ができるのでは？

○研究データは分野により性質があまりに異なるため、部局レベルの運用ガイドラインで、論文および根拠データ（実験ノートまで含めるのであれば根拠資料）は、機関リポジトリで管理、とすることは可能か？

← 現在の機関リポジトリシステムでは、データ保全機能（凍結・非改変・データのバージョン管理・タイムスタンプ等）が不十分か？

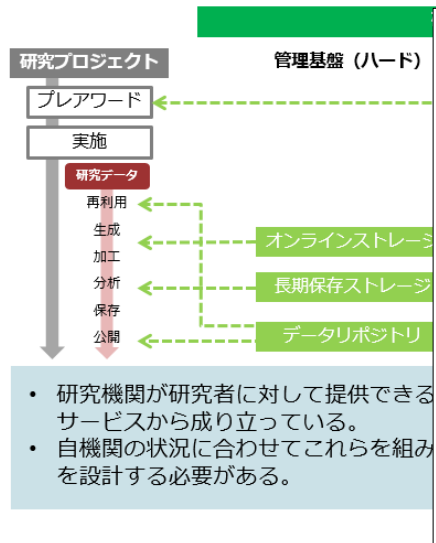
J P C O A R

オープンアクセスリポジトリ推進協会

第7章 研究データ管理サービスの設計

- ・ 自機関で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

7.1.1 研究データ管理サービスとは？



7.2.4 サービスの設計：調査 (1)

- ・ サービス設計の前にまずは調査の実態やニーズを把握する

目的

- ・ 研究者が取り扱っているデータの実態を把握する
- ・ 研究者の研究データ管理に対する予備計画に活かす。

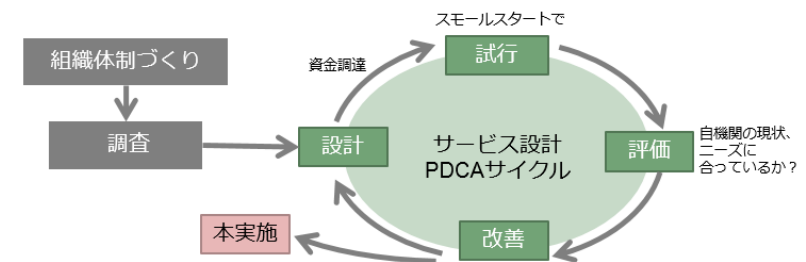
推奨される方法

- ・ アンケート調査
- ・ インタビュー調査

参考

- ・ DAF (Data Asset Framework) = 研究データで実施例がある。
<http://www.data-audit.eu/>
<http://www.data-audit.eu/docs/DAF-Imp.pdf>
- ・ RDM Rose 2015 > 3.1 DAF surveys / 2
<http://rdmrose.group.shef.ac.uk/?page=2>

7.2.5 サービスの設計：試行、評価



- ・ 研究データ管理サービスの定着は長期的計画で
- ・ オンラインストレージなど管理基盤の提供（ハード）、相談窓口やトレーニングなどの人的支援（ソフト）ともに、試行と評価、改善をくりかえしながら、自分の機関に合ったサービスを構築し、組織内の意識を高めていくことが重要

今後の展開

- RDMトレーニングツールの活用促進
 - RDMトレーニングツールのMOOC配信
 - JMOOC/gaccoでの開講
(2017.11.15開講)
 - セミナーの開催
 - 情報リテラシー教育への組み込み
- 持続的な教材開発体制の確立
 - 内容のアップデート
 - 教材の段階的な拡充と高度化

JMOOC/gaccoで開講

<http://gacco.org/>

Research Data Management

2017 11/15 開講

無料オンライン講座、今秋開講

オープンサイエンス時代の研究データ管理

目的

- 研究者の研究データ管理 (RDM: Research Data Management) に関する基礎的な知識を習得する
- 各機関における RDM サービス構築の促進を図る

講座ページ

https://nii.gacco.org/course/course-v1:gacco+ga008+2017_11/about
お申し込みは 8/23 に公開予定です。

内容

- 第1講 研究データ管理とは
- 第2講 保存・共有・文書化
- 第3講 メタデータ・知識発見の関わり
- 第4講 講義に向けて

受講対象者

- 図書館、IT 部門、研究支援部門の職員 (IRAS) 等
- オープンサイエンスや研究データ管理に関心のある教職員

8/23 から
お申し込み
受付開始!

QRコード

nii.gacco.org

協賛: JPCOAR
実行: NII

Research Data Management

はじめての研究データ管理!

OPEN SCIENCE

Research Data Management Training Tool

「はじめての研究データ管理！」

**リサーチデータマネジメント
トレーニングツール**

オープンアクセスリポジトリ推進協会 (JPCOAR) 研究データタスクフォースでは、2016 年度に研究データ管理のための教材「RDM トレーニングツール」を作成しました。スライドやスクリプトを公開しています。この教材が元になり、本講座が生まれました。

QRコード

RDM トレーニングツール

<http://id.nii.ac.jp/1458/000000023/>

オープンサイエンス時代の到来に向けて準備を進めている、国立情報学研究所の新しい学術情報基盤について、講演動画と講演録を公開しています。

QRコード

オープンサイエンスを支える新しい学術情報基盤

山崎 一樹 NII オープンサイエンス基盤研究センター センター長
<http://www.nii.ac.jp/openhouse/report/#ohkey>

NII オープンサイエンス基盤研究センター

研究データの管理・公開・利用のための共通のシステム基盤を提供します。

研究スタイルのパラダイムシフトとなるオープンサイエンスの推進に
内けて、研究データの保存・公開・検索のための基盤を創設・維持のもとで
研究開発し、国内の大学や研究機関とともに利便性の促進を図ります。

NII 国立情報学研究所
National Institute of Informatics



受講をお願いします！