eScienceへの広がり: Open Repository

DRF, 岡山大学附属図書館 大園 隼彦



本日のテーマ

- Open Repositories 2012の概要
- Research Data Management
- 国内のこれからの展開を考える



Open Repositoriesとは

- デジタルリポジトリの開発や運用に携わる関係者が世界中から集まり、情報の共有や議論を行う国際会議。
- Open Repositories 2012

7回目の開催。

2012年7月9-13日にエジンバラ大学で。

30カ国超から450人以上の関係者が出席。



OR2012参加報告のダイジェスト①

- 共同リポジトリ・共同インフラOAを促進するための共同リポジトリ・共同インフラ
 - *UK RepositoryNet+:
 UKの助成研究の成果を公開する共用インフラを用意。
 - OpenAIRE: EUの助成研究の成果を公開するためのインフラ整備とリポジトリの連携促進によるアクセス性の改善

参加機関の増加=コンテンツの量の増加 サービス連携等によるコンテンツのアクセス性の改善



OR2012参加報告のダイジェスト②

• 研究データ管理

目的:データの長期保存、データの永続アクセス、 データの再利用、データの引用 等

- 研究データを保存し、統合検索するためのインフラを整備。
 - Research Hub (Griffith Univ.)
 - DataFinder (Oxford Univ.)
 - Open Exeter (Univ. Exeter)
 - ICAT (ISIS)

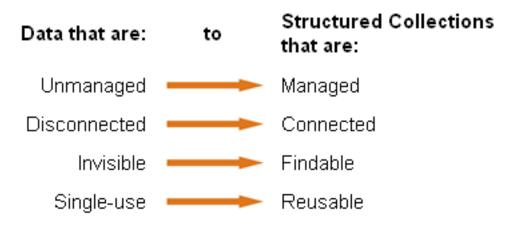




Australian National Data Service

Our Vision: More Australian researchers reusing research data more often

ANDS is enabling the transformation of:



http://ands.org.au/



研究データ管理に関する最近の情報

• トムソンロイター、自然・社会・人文科学分野の様々なデータセットを検索可能な"Data Citation Index"をリリース

(カレントアウェアネス 2012/10/17)

• 米カリフォルニア電子図書館等、表形式の研究データを管理・公開するWebアプリケーション"DataUp"をリリース

(カレントアウェアネス 2012/10/3)

- ANDSの研究データのメタデータがPrimoCentralのインデクスに収録 (カレントアウェアネス 2012/09/13)
- Journal of Heredity誌、論文中で使用されたデータの公開を求める"Joint Data Archiving Policy"を採用へ

(カレントアウェアネス 2012/07/23)

 BioMed Central、研究データを公開しジャーナル論文にリンク (STI Updates 2012/4/6)



日本でも研究データ管理を 始めるべきではないか?

- 研究データ管理に図書館のノウハウが生かせるはず。 データの長期保存、データの永続アクセス、 データの再利用、データの引用
- 目的: まず保存すること、次に引用できる環境を。
- 考えること
 - 何を保存するのか
 - どこに保存するのか?
 - どのように保存するのか?
 - サービスの展開はどうするのか?



研究データ管理 ガイドライン 一OR2012 Monash大学より一

- データプラットフォームの選択 研究コミュニティが既にプラットフォームを持っている場合は、 それを採用し支援する。
- データプラットフォームの開発 開発は最終手段。
 特定分野のニーズにマッチし、他機関でも利用できるものを開発。
- データプラットフォームの展開 メタデータの充実、識別子の付与、データ連携。

(Institutional Infrastructure for Research Data Management https://www.conftool.net/or2012/index.php?page=browseSessions&form_session=12)



何を保存するのか

- 研究データ = 研究で利用する全ての情報
 - 構造化データ
 - : データベース、表等
 - ・非構造化データ
 - : テキスト、イメージ、音声、記録、メール 等

(Research data management – an introduction DaMaRO project Oct. 10, 2012 http://damaro.oucs.ox.ac.uk/induction.xml)



どこに保存するのか

- 既存のプラットフォームの活用図書館でサポートするならば機関リポジトリを活用する相互運用性での貢献
- その他の可能性
 - ○研究者はコミュニティのシステムをより利用する傾向にある。 例) Dryad: 生物科学分野のデータリポジトリ
 - 〇新設のプラットフォーム 最終手段、リスクが高い。 例) Research Hub、DataFinder、OpenExeter、ISIS



国内の状況(機関リポジトリ)



すべてにチェック すべてのチェックをはずす 資料の種別を指定する 学術雑誌論文 (311,910 件) 学位論文 (77,007 件) 会議発表論文 (87,672 件) 紀要論文 (602,996 件) 図書 (22,311 件) 会議発表用資料(11,787 件) <u>テクニカルレポート (16,094 件)</u> 研究報告書 (26,135 件) 一般雑誌記事 (56,212 件) <u> ブレブリント(354 件)</u> データ・データベース (52,659 件) 教材 (8,419 件) <u>ソプトウェア(31 件)</u> その他 (154,877 件)



国内の状況(機関リポジトリ)



検索対象資料種別:データ・データベース

検索対象機関: 千葉大学

検索結果 52223 件中 1 ~ 10件を表示

10件ずつ表示 ▼

該当レコードが1000件を超えているためソートできません。



"子供のための家具" pour des enfants 作業机(タイプ-D) / 小山 利明

空間工房Gen木

No Image 千葉大学学術成果リポジトリ 千葉大学

「単1後日本デザインの軌跡 1953-2005 千葉からの挑戦11展会期:2006年4月1日-5月28日...

"子供のための家具" pour des enfants 小椅子(タイブ-E) / 小山、利明

空間工房Gen木

No Image 千葉大学学術成果リポジトリ 千葉大学

「単5後日本デザインの軌跡 1953-2005 千葉からの挑戦10人会期:2006年4月1日-5月28日...



どのように保存するのか

- メタデータの充実 OAジャーナルのようにフルテキストの検索は期待できない。 充実したメタデータがアクセス性を改善する。
- 識別子の付与
 データセットに永続IDを付与することで、その引用やアクセス
 性が改善される



メタデータの充実(参考)

- DataCite Metadata Schema v.2.2
 - ★必須項目:

識別子、作成者、タイトル、公開者、公開年

☆任意項目:

分類・キーワード、寄与者、日付、言語、資源タイプ、

その他のID、サイズ、フォーマット、バージョン、権利、

その他詳細情報

☆管理情報:

最終更新日、メタデータバージョン

〇 相互運用性を考える

(http://schema.datacite.org/meta/kernel-2.2/doc/DataCite-MetadataKernel v2.2.pdf)



識別子の付与

DOI

"Research Hub"はANDSがDOIを付与

"DataFinder"ではDataCiteがDOIを付与予定

国内: 今後の"JaLC"の展開に期待

• 研究者ID

Research Hub、ANDS、NLA (National Library of Australia) は相互にリンク

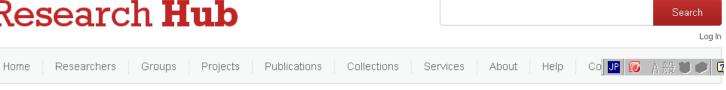
国内: 金沢大学CSIプロジェクト

科研費番号やRNR-IDにより著者識別(名寄せ)

ORCIDの動向も視野に



Research Hub



Research Hub » Collections » Braided Channels

Braided Channels (2011)

Short Description

The Braided Channels Research Collection includes materials collected on Australia women, land and history in the Channel country. The collection is constructed from some 70 hours of oral histo (... more)

Digital Object Identifier	<u>10.4225/01/4F8E1281B8E2A</u>
Location	Griffith Film School, Griffith University, South Brisbane Qld 4101
Managed By	Griffith Film School Trish Fitzsimons
Output of	The Braided Channels creative research project - Part Two











サービスの展開はどうするのか

- まずは保存し、次に利用の環境を提供する。
- より多くのデータが保存され、さらに再利用されるためにポリシー の作成が必要。
- 他システムとの相互運用性
- データ再利用の仕組み



参考



研究データ管理支援を開始する図書館への10の提言 - LIBER (欧州研究図書館協会)-

- ① 研究データ管理の支援活動(助成金申請、知的財産の助言等)を 提供する。
- ② メタデータ及びデータの基準の発展に貢献し、メタデータサービスを提供する。
- ③ データライブラリアン職を作り、専門職員を育成する。
- ④ 組織のデータポリシー開発に積極的に参加する。データのライフサイクルに適したオープンデータポリシーの採用を促す。
- ⑤ 研究者やデータセンターと連携し、データの相互運用基盤(共有、アクセス性)を促進。



研究データ管理支援を開始する図書館への10の提言 - LIBER (欧州研究図書館協会)-

- ⑥ 保存、発見、永続アクセスを提供することで、研究データライフサイクル支援をする。
- ⑦データに永続IDを付与し、データの引用を促進する。
- ⑧ 既存のインフラを活用して保存、検索サービスを提供する。
- ⑨ 特定分野のデータ管理の実践に参加する
- ① IT部門との協力、クラウドサービスを活用することで、動的・静的研究データの安全な保存を実現する。

(http://current.ndl.go.jp/node/21693)



研究データ管理ポリシー University of Edinburgh

- 1. 研究データは研究データライフサイクルを通して最高基準で管理 される。
- 2. プロジェクトにおける健全な研究データ管理の責任は研究責任者にある。
- 3. 全ての新しい研究企画は、データの保存、管理、完全性、信頼性、 共有、公開に取り組む、研究データ管理計画を含まなければなら ない。
- 4. 大学は研究データ管理計画のガイドラインやトレーニング、サポート、アドバイスを提供する。
- 5. 大学は、今後の研究データのアクセスのため、その保存、バックアップ、登録のサービスとシステムを提供する。



研究データ管理ポリシー University of Edinburgh

- 6. 他の場所で管理されているデータ、例えば、国際的なデータサービスや サブジェクトリポジトリ、は大学にも登録する。
- 7. 研究データ管理計画は、適切な安全対策の下でデータがアクセス、再利 用されることを保証しなければならない。
- 8. 分野研究データの合法的な所有は守らなければならない
- 9. 今後歴史的に重要になる研究データ、研究成果を検証するデータを含む 全ての研究データは、適切な国または国際的データサービス、分野リポ ジトリ、大学リポジトリでの登録、保存を提供され、評価される。
- 10. ファンドの条件に関わらず、データをオープンに再利用する権利を保持 せずに、データを公開・再利用する排他的な権利を商業出版社に渡すべ きではない。

(http://www.ed.ac.uk/schools-departments/information-services/about/policies-and-regulations/research-data-policy)



まとめ

- データ管理に既存のシステムを活かす
 - → 機関リポジトリを利用できる
- 新しくデータ管理システムを構築する
 - → 機関リポジトリ構築のノウハウを活かして、システム連携に貢献する

