



Title	月刊DRF 第72号
Author(s)	デジタルリポジトリ連合
Issue Date	2016-01-12
Doc URL	http://hdl.handle.net/2115/73639
Type	periodical
Note	事務局: 北海道大学附属図書館; http://drf.lib.hokudai.ac.jp/ で公開したもの
File Information	DRFmonthly_72.pdf



[Instructions for use](#)



月刊 DRF

Digital Repository Federation Monthly

第72号
No. 72 January, 2015

【特集1】DRF 研究データに関するオンラインワークショップを開催中
 【特集2】機関リポジトリ担当者のための（もう一度）OAI-PMHって何？
 【連載】いまそこにあるオープンアクセス 第16回
 「学位論文をプロクエストに提出するよう義務付けているか？」

【特集1】DRF 研究データに関するオンラインワークショップを開催中

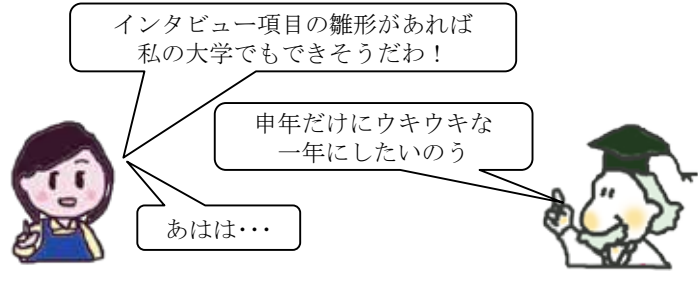
DRFでは昨年11月18日～2月末まで機関リポジトリ担当者向けにオンラインワークショップ「研究データから研究プロセスを知る」を開催しています。

オープンサイエンスの潮流が高まる中、機関リポジトリでも研究データを扱う必要が今後高まっています。しかし、ほとんどの機関リポジトリ担当者は研究者が研究の過程でどのようなデータを扱っているのかよく知りません。今回のワークショップでは、様々な分野の研究者に研究プロセスの各段階でどのようなデータをあつめているのか詳細に聞き取り、分野ごとに事例集を作成することで、今後研究データを扱う際の参考資料とすることを目的としています。

15名の参加者が担当する研究分野ごとに複数の班に分かれ、班内でインタビュー項目を作成し、それに基づき各々が研究者にインタビューを行います。インタビュー項目の作成にあたってはパデュー大学の

のポスター：“Conducting a Data Interview”（DRFで和訳）を参考にしてしています（下図）。このポスターは図書館職員が研究者にデータインタビューを行う際のとっかかりとして利用できる10の質問群を紹介したものです。

現在は各班でインタビュー項目の作成が終わり、研究者にインタビューを行っています。本ワークショップで作成したインタビュー項目やインタビュー結果等の成果物は後日DRFウェブサイトで公開する予定です。



Conducting a Data Interview

Michael Witt & Jake Carlson, Purdue University Libraries.

10の質問群

1. そのデータのなりたち
2. そのデータはどんな書式・形式か
3. そのデータセットの寿命はどの程度と予想されるか
4. そのデータはどのように利用・再利用・再定義され得るか
5. そのデータはどの程度の大きさで増加率はどの程度か
6. そのデータの潜在的な読者は誰か
7. そのデータはだれが所有しているか
8. そのデータセットは取扱いに注意な情報を含んでいるか
9. そのデータからどんな文献や発見が生じるか
10. そのデータはどのように公開されるべきか

(翻訳の全文)

http://drf.lib.hokudai.ac.jp/drf/index.php?plugin=attach&refer=Foreign%20Documents&openfile=purdue_poster_trans.pdf

“Conducting a Data Interview”
Michael Witt & Jake Carlson, Purdue University Libraries, West Lafayette, Indiana, USA

もう一度 OAI-PMHってなに？



機関リポジトリでコンテンツを公開すると、さまざまなサービスでも自動で検索できるようになりますね。例えば雑誌論文ならCiNiiで検索できるようになってとっても便利です！

そうじゃな。そのような連携を可能にしている、OAI-PMHという基本的な仕組みがあるのじゃよ。



OAI-PMHですか？なんだかよく聞くのですが、あまりしっかり考えたことがなかったです。

それでは、一年の最初でもあるから、今回は基本に立ち返りOAI-PMHについて確認してみようかの。



OAI-PMHとメタデータ

OAI-PMHは、Open Archives Initiative Protocol for Metadata Harvestingの略称で、Open Archives Initiative¹⁾という団体が作った、メタデータをハーベスト（刈り取り）するための通信上の決まりごとです。

機関リポジトリでは、コンテンツを登録する際に、コンテンツ毎にその書誌情報を付与します。例えば雑誌論文なら、論文名、著者、雑誌名、巻号、掲載ページなどの書誌情報が必要となります。この書誌情報のことをメタデータと呼んでいます。メタデータは機関リポジトリ内にコンテンツと結び付けて保存されています。

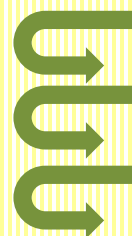
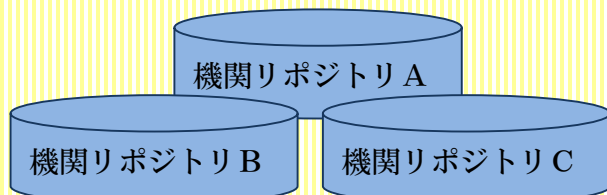
機関リポジトリ担当者の図書館職員にとって、書誌情報と聞いてすぐに思い浮かべるのが、CiNii Booksで公開されているNACSIS-CATの図書・雑誌情報ではないでしょうか。NACSIS-CATは国立情報学研究所（以下NII）が運営している目録データベースで、各図書館はNIIが定めたルールに則って、NIIのNACSIS-CATに書誌や所蔵を登録します。

一方、機関リポジトリではメタデータは各機関のシステムに登録されますので、メタデータ入力のためのルールなども各機関で決めることになるのです。

OAI-PMHの役割

NACSIS-CATなどと違い、機関リポジトリのメタデータは機関リポジトリごとに保存されています。したがって、そのままでは複数の機関リポジトリのメタデータをいっしょに検索することができません。そのような情報をまとめて検索するには、各機関リポジトリのメタデータを収集してあげる必要があります。そのとき重要な役割を果たすのが、OAI-PMHなのです。

この収集の方法が、あたかも作物を収穫するようになるため、メタデータのハーベスト（刈り取り）と呼ばれています。刈り取られる側のシステムをリポジトリと呼ぶのに対し、刈り取る側のシステムをハーベスタと呼ぶこともあります。国内の機関リポジトリにとって一番身近なハーベスタはNIIのIRDB²⁾になるでしょう。



OAI-PMHの動作

OAI-PMHによるシステム間の通信は、一般のwebサイトと同じhttpという方法で行われます。機関リポジトリの各ページは人が閲覧するために、htmlで記述されています。webブラウザでアクセスすることによってhttp通信が行われ、人が見やすいように整形され閲覧できるようになっています。OAI-PMHではシステム間の通信のための専用のページにあたるものが、システムにとって扱い易いXMLという形式で用意されています。この専用ページのアドレスをBaseURLと呼びます。

BaseURLの例: <http://ir.abcd.ac.jp/oai/request>

すべての機関リポジトリはOAI-PMHに対応していますから、BaseURLもすべての機関リポジトリが持っていることとなります。言い換えれば、BaseURLを公開していない機関リポジトリはリポジトリとは呼べないのです。

ハーベストの際には、どのようにデータをハーベストしたいのかをverbという値で指定

することができます。verbには6種類の値がありますが、全てのコンテンツのメタデータを収集するにはListRecordsを指定します。また、ListRecordsを行う際にはメタデータの形式をmetadataPrefixという値で指定します。すべての機関リポジトリは国際的なoai_dcというダブリンコアの形式に対応していますし、日本国内であれば、NIIが策定したjunii2という形式にも対応しているはずです。

ハーベストの指定例: (実際には改行なし)

```
http://ir.abcd.ac.jp/oai/request
?verb=ListRecords
&metadataPrefix=junii2
```

上記の例では、この機関リポジトリ全てのコンテンツのメタデータをjunii2という形式でハーベストすることができます。一般のwebブラウザでも表示できますので、一度自分の機関のリポジトリのBaseURLを指定してアクセスしてみてください。

OAI-PMHの使われ方

NIIのIRDBによってハーベストされたメタデータは、NIIのJAIRO³⁾というサービスで検索できるようになります。機関リポジトリごとに分散していたメタデータが一度に検索できるようになるのです。収集されたメタデータはさらに別のシステムにハーベストされることもあります。IRDBのメタデータはジャパンリンクセンター(JaLC)にハーベストされることによりDOIが付与されたり、国立国会図書館にハーベストされることにより、博士論文の全文ファイルを収集すること

が可能になっています。このように、OAI-PMHは機関リポジトリのサービスをより高めるために使われています。

また、OAI-PMHはメタデータを一括で収集することに優れているため、機関リポジトリ以外にも使われています。例えば、図書館などが提供する「ディスカバリーサービス」では、さまざまな情報源のデータを一括して検索できますが、そのデータの収集にはOAI-PMHが使われているのです。

¹⁾ Open Archives Initiative
<http://www.openarchives.org/>

²⁾ IRDBコンテンツ分析
<http://irdb.nii.ac.jp/analysis/index.php>

³⁾ JAIRO
<http://ju.nii.ac.jp/>

ところで、OAI-PMHでハーベスト(刈り取り)されたからといって、機関リポジトリのメタデータが取られてしまうことにはならないから安心したまえ。要するにコピーを取るのと同じことだと考えればよいのじゃ。



学位論文をプロクエストに提出するよう義務付けているか？

Clear and present open access 16. Are students required to submit their dissertations and theses to Proquest?

[本誌2015年1月号](#)で、欧米における学位論文のリポジトリ登録義務化をめぐる議論についてお伝えしたが、およそ1年が経過した11月下旬、同じ[SCHOLCOMM](#)メーリングリスト上で、再び学位論文の扱いに関する議論が開始された。

発端はロヨラ大学の図書館員からの、大学が学位論文の提出要領を見直しており、院生にプロクエスト([ProQuest](#))への提出を必須とするかどうかの問題になっているので、他大学の事情が知りたい、という[質問](#)だった。プロクエストは、ご存知の方も多いと思われるが、米国議会図書館からも認定されている学位論文データベースの出版社である。前身のUMI社の時代から全米の多くの大学と提携し、博士論文のマイクロ化、複製の販売などを行っており、米国の学位論文の流通に関して長らく独占的な地位を占めて来た。

これに対し、モンタナ州立大学の図書館員から、博士論文のみ必須になっているが、電子版学位論文(ETD)がリポジトリに登録されているのでメリットがなく、義務付けをやめてほしいと大学院側に要望しているという[回答](#)があった。ジョンズ・ホプキンス大学もETD化をきっかけにプロクエストへの提出義務付けはやめた[とのこと](#)である。

一方、ケント州立大学(オハイオ州)では[FAQ](#)でETDについて学生向けに詳しく説明している。それによれば、大学に提出されたETDは、[オハイオリンクETDセンター](#)という組織で管理されると同時に自動的にプロクエストへも提出される。間接的に必須となっているわけである。あなたの論文がプロクエストから売れば印税を受け取れるかもしれないなどとも書かれている(オハイオリンクETDセンターから同じものが無料公開されているのだが)。

派生して、学生からリポジトリ公開の許諾

を取っているかという[質問](#)も出て、デューク大学ではリポジトリでの公開が学位取得の条件となっている(エンバーゴ申請は可能)との[回答](#)があった。また、カーネギー・メロン大学の図書館員から、リポジトリへの登録は強く推奨されているが必須ではなく、プロクエストへの登録もオプション、すなわち同大図書館では博士論文の網羅的収集はできていないとの[告白](#)などもあった。

実はプロクエストに関する各大学の方針をまとめた簡単な[リスト](#)はすでに作成・公表されており、この一連の議論の中でも真っ先に紹介された(ただし、このブログはすでに更新を停止しており、情報が古い)。それとは別の[リスト](#)もグーグルのスプレッドシート(表計算と訳すと何か違う)で公開されている。

メーリングリストに寄せられた情報の[まとめ](#)を見ると、有力大学のプロクエスト離れが起こっているように感じられる。プロクエストはETDの国際的な連合組織[NDLTD](#)のパートナーでもあるのだが、各大学がリポジトリで学位論文を提供する方向に進むとしたら今後どうなるのか。一商業出版社のことながら気になるのである。

栗山 正光
(首都大学東京
学術情報基盤センター
教授・DRFアドバイザー)



【researchmap】

<http://researchmap.jp/read0195462>

次号
予告

【特集】 オープンアクセス・ポリシー事情

【連載】 かたつむりとオープンアクセスの日常