

機関リポジトリ概論

平成 18 年度 学術ポータル担当者研修

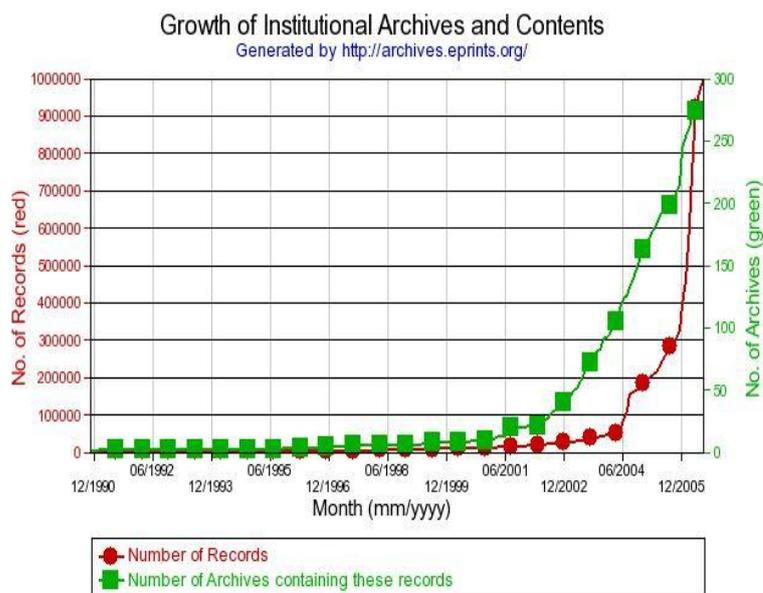
村上 祐子（国立情報学研究所）

キーワード：機関リポジトリ、学術コンテンツ、オープンアクセス

機関リポジトリの現状

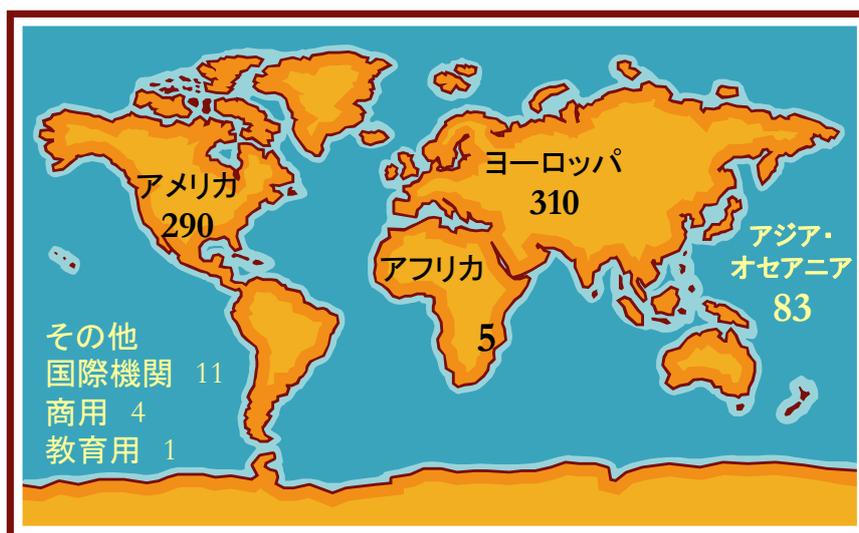
プレプリントサーバを機関リポジトリと呼ぶかどうかには議論の余地があるが、少なくともデジタルリポジトリであることは間違いない。世界初のプレプリントサーバ arXiv (物理学) は 1991 年に立ち上がった。その後 1995 年までに、経済学 (WoPEc, その後 RePEc に継承)・核物理学 (CERN)・計算機科学 (DARPA)・宇宙科学 (NTRS) などの分野リポジトリが運用を開始した。機関リポジトリとしては、1996 年にバージニア工科大学が Networked Digital Library of Theses and Dissertations を創始したのがはじめてである。

その後もじわじわと機関リポジトリは増加した。下のグラフを見ると、2002 年に入ってから急激にリポジトリ数が増加し、2005 年になってコンテンツ数が飛躍的に伸びたことがわかる。



2006 年になると、1-2 日に 1 つの割合でリポジトリが新しく運用を開始している。

世界の機関リポジトリ:704 (2006年6月末)



<http://archives.eprints.org/>

機関リポジトリの分布は上図のようになっている。大学の役割は国によって違い、単純には普及状況を比較できないが、大学数を勘案すると、ヨーロッパの大学は機関リポジトリがある割合が他の地域に比べて多い。アメリカでは、2年制大学を含めるとリポジトリ所持率は6%程度、4年制大学(2,530大学)に限れば11%となる。一方、ヨーロッパは全体でも700大学程度なので4割強である。特に、ドイツでは1/5、イギリスでは1/3、オランダではすべての大学が、すでにリポジトリを備えている。すなわち、世界のトップクラスの研究大学は知的資産の管理の重要性を認識し、リポジトリを備えているというのが現状といえる。

日本では4年制大学は約700校で、2006年6月現在19のリポジトリが運用されている。

機関リポジトリの定義

現実と理念の差に注意しながらリポジトリの定義を眺めていこう。

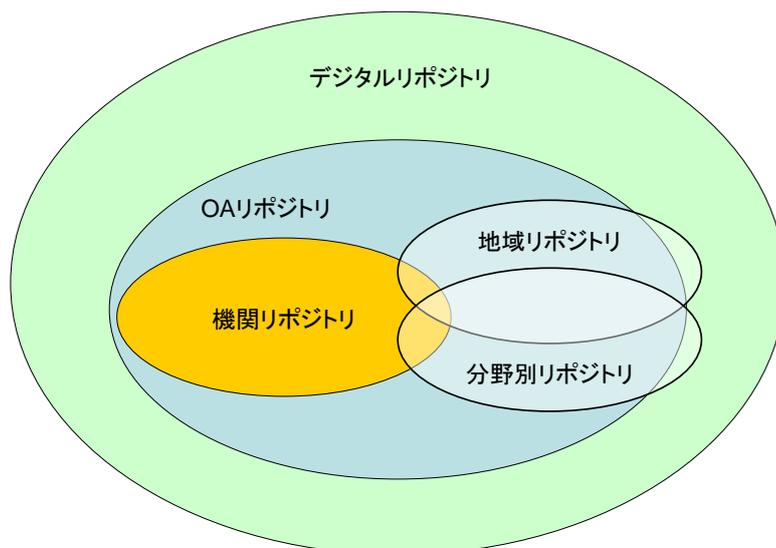
機関リポジトリは、デジタルリポジトリのうち、大学等学術機関が学術目的で管理運営するものである。また、本来学術コンテンツはオープンアクセスで

あるという理念的原則によれば、デジタルリポジトリの部分集合であるオープンアクセスリポジトリのうち学術機関が管理運営するものということもできる。

機関リポジトリには、コレクションとしての側面とサービスとしての側面があり、次の二つの定義はそれぞれに重点をおいたものである。

1. **リンチの定義**：大学がそのコミュニティのメンバーに提供する、大学およびそのコミュニティのメンバーにより創造されたデジタル資料の管理と配布のための一連のサービス
2. **クロウの定義**：大学の知的成果を蓄積保存するデジタルコレクションで、機関で範囲限定（専門分野別ではない）・学術的・累積的かつ永続的・オープンで相互運用可能なもの

デジタルリポジトリと機関リポジトリ



クロウの定義によれば機関リポジトリと分野別リポジトリは重ならないこととなるが、現実には重なっているケースもある。

デジタル資料のコレクションという側面に注目した場合、機関リポジトリと従来の電子図書館との違いが問題となる。国立国会図書館(2002)によると、電子図書館の特徴としては次の条件が挙げられる。

- ・ ネットワークによる情報の提供
- ・ 資料にアクセスさせるための書誌情報（二次情報）、電子化した資料その

もの（一次情報）など、さまざまな電子図書館の「蔵書」（コンテンツ）の構築と提供

- ・ テキストだけではない音声、動画などマルチメディアの活用
- ・ 情報通信技術を活用した検索・閲覧等の利便性の向上
- ・ インターネット上の情報など外部情報資源の活用と探索への援助

これをクロウの定義と比べてみると、電子図書館でもクロウの条件を満たしていれば、リポジトリとみなすことができることになる。だが現状の電子図書館ではそれが実現されているとはいえず、大学が管理運営している電子図書館をリポジトリとするためには、コレクションの見直しおよびサービスの再構成が必要となる。たとえば、第3の条件「累積的」を満たすためには、コレクションが「生きて」いなければならない。つまり、貴重書を電子化しただけでは、リポジトリとはいえない。また、最後の条件「オープンで相互運用可能」は本質的である。というのも、リポジトリの最大の特徴は、「利用者は、わざわざそのコレクション本体にアクセスしなくても、さらにそのコレクションの存在・設置場所すら知らずに、横断検索等の手段を用いて、求める資料を利用することができる」という点にあるからである。そのため、たとえ資料が公開されていたとしても単にウェブで発信するだけの電子図書館はリポジトリとは認めがたい。相互運用可能条件を満たすためには、実装上最低限の条件として、標準プロトコルを用いたメタデータのハーベスティングが可能でなければならない。

学術コンテンツ

学術コンテンツとは、

- 学術誌論文
- 学位論文
- 紀要論文
- 報告書
- 電子ブック
- 特殊コレクション：貴重書・博物資料
- データベース
- 教材：講義ビデオ・シラバス
- ファクトデータ：実験・観測データ
- 画像：設計図・写真

などであり、しかもこれらに限ったものではない。いわば教育研究に有益な電子的資料であれば、「なんでもあり」であるといえる。

学術コンテンツとオープンアクセス

オープンアクセス (open access, OA) とは、研究成果を研究コミュニティ全体にオンラインで自由に (無料で) 利用可能にする手段である。インターネットは、どこでもだれでもいつでも、資料を利用可能にするための手段のひとつである。これは、学会発表論文や学位論文、研究レポートのような他のタイプの研究成果と同様、学術誌論文にも当てはまり、そしてそれはその論文が発表されている学術誌を自分たちの図書館が購読しているかどうかに関わりがない。その意味で OA モデルは既存の購読出版モデルに対立するものといえる。

OA 思想では、学術的ならば OA であるという原則がある。いいかえると、OA でないものは、学術的ではない。つけくわえると、学術的でなくても OA であるものがある。直接の金銭的収入が期待される成果には OA は適用されない。つまり、成果から直接の金銭的収入が期待されるならば学術的ではない。したがって、この原則を認めるならば、たとえば特許収入が期待される成果は純粋には学術的なものとはいえないことになる。

OA 文献とは、どこでも誰でも自由に利用可能な文献のことである。学術的なものが OA であるならば、学術文献は OA 文献である。文献を OA で提供する手段としては、OA ジャーナルと OA リポジトリがある。OA リポジトリ (Green road) とは、蓄積保存した資料を世界中に無料で利用可能にするである。オープン・アーカイブ・イニシアチブ (Open Archive Initiative: OAI) のメタデータ・ハーベスティング・プロトコルに対応していれば、他のアーカイブと相互運用が可能となる。

一方、OA ジャーナル (Gold road) では、査読を行い、受理された論文を世界中に無料で利用可能にする方式で、読者が自由に学術論文にアクセスできる電子ジャーナルであり、一般的な学術雑誌のスタイルを踏襲したものである。出版者側は既存の手法にこだわりがちであるが、OA と適合したビジネスモデルを作れるかどうか、つまり査読、編集、システムにかかる出版費用をどう負担するかが問題となる。大まかに分けて次の3つの方式がある。

1. 「同人誌」型：編集業務を研究者自身が無償 (あるいはごく低報酬) でおこなう。
2. OA ジャーナル型 (アメリカ物理学会など)：出版費を著者または所属機関が投稿料として負担することとなる。だが査読費用は採択率に依存し、雑誌出版費、ひいては投稿料の算出根拠は一般的には難しい。(倉田(2005))
3. OAハイブリッド：著者は従来の購読モデルでの出版とOA投稿料モデルでの出版を選択することができる。この方式ならば、研究助成団体からOAでの成

果公開を要請されたとしても、商業出版者によるジャーナルに投稿することができるようになるため、2006年になって、商業出版者でもOAハイブリッドジャーナルを開始するようになった。オープンアクセスジャパンによるまとめは次表の通りである。

(http://www.openaccessjapan.com/archives/2006/05/post_58.html)

出版者	名称	対象タイトル数	価格(\$)	著作権の扱い
Springer	OpenChoice	全タイトル?	3000	CC
Blackwell	OnlineOpen	79	2500	-
OUP	Oxford Open	47	1500~2800	通常と同じ
Elsevier	なし	6	3000	出版者へ委譲
日本化学会	Open Access Option	2	500~1000	-

だが現実には、OA ジャーナルだけではなく、OA ハイブリッドジャーナルの実現も難しい。たとえば、日本の大学所属研究者の一人当たりの研究費はそれほど多くない。

投稿料の許容額	パーセント
払わない	11
500 ドルまで	45
1000 ドルまで	18
1500 ドルまで	5
2500 ドルまで	2
わからない	17
無回答	2

SPARC-JAPAN のアンケートの結果はこの現実を反映しており、日本の大学所属研究者の投稿料に関する意識調査は上の表の通りである

(<http://www.nii.ac.jp/sparc/shiryou/index.html>)。2つの表を比べると、日本の研究者の研究費ではOAハイブリッド誌に投稿しない・できないこととなる。

このように、研究助成機関や所属機関によって研究成果のOA公開が要請され、かつ研究者が投稿料モデルのOAジャーナルに投稿しない・できないならば、OAの実現方式としてはリポジトリが有望となる。この論点と、学術的ならばOAでなければならないというOA思想の原則、さらに大学の社会責任論（情報公開）という論点が結びついて、機関リポジトリが要請されることとなる。

コンテンツによる機関リポジトリの特徴づけ

しばしば、コンテンツによって機関リポジトリを特徴付けることがある。まず、学術論文型リポジトリでは、学術論文を中心的に蓄積・発信する。また、「電子図書館型」というより学術データ型リポジトリでは、図書・雑誌論文だけでなく、研究活動に有益なさまざまな電子データを蓄積・発信する。

オープンアクセスの原則から見ると、論文、とりわけ学術誌論文を自由に利用可能にするための実装手段のひとつがりポジトリである。つまり、学術論文を重視しないデジタルコレクションはりポジトリではない。この論点を押しすすめて、電子化した貴重書等論文以外のデータが入っているデジタルコレクションはりポジトリとしては認めがたいとする議論は、商業出版者による学術誌出版のモデルとOAモデルの対立が中心的関心事となるOAの議論の影響としてはもっともである。だが、その結果、広義の研究成果であるプレプリント・研究会資料・データなど、研究活動の現場では価値が高く不可欠なものである学術コンテンツが、素性のはっきりしないグレイ・マテリアルとして軽視されることとなりかねないようでは、研究支援という本来の学術情報流通の目的からすれば本末転倒である。

実際、2005年末に韓国でのES細胞研究をめぐる疑惑が発覚、捏造と断定されたのを契機に、2006年に入ってから研究に用いられるデータの公開を要請する動きが研究資金助成団体からも生じることとなり、オープンアクセスの実現を支持するSPARCもオープン・データの調査に乗り出すこととなった。

OA思想がオープン・データも包摂していくなれば、今後はこのような広義の学術コンテンツを対象とする機関リポジトリが中心となっていくと考えられる。だが現状では、運用が難しく、今後の課題である。SPARCの展開、また国内外の研究助成機関・研究者の動向に注目する必要がある。

機関リポジトリの直面する課題

機関リポジトリの運用における課題は大きく分けて、立ち上げ期に見られる課題、そして、安定運用上の課題がある。

立ち上げ期には、全学的合意形成、システム選定・構築・運用体制確立、コンテンツ確保・登録ポリシー形成を平行して行う必要がある。

また運用上の課題として不可避なものは、著作権の問題である。著者が著作権を保持している場合もあるが、出版契約時に出版者に著作権の権利譲渡を行うことがある。この場合には、出版者が権利者としてさまざまな許諾を与えることになる。もっとも実際には、電子ジャーナルが一般化した1998年以前には、出版社と著者との間で交わされた契約書にはまったく電子的学術情報に関する許諾条項はないという指摘がある(Feijen 2006)。それが正しければ、著者が承認する限りでは、1998年以前にさかのぼって過去の教育・研究成果を機関リポジトリに登録・発信することは著作権の観点からはまったく問題とならないことになる。1998年以降の契約では、現在では出版社や学協会が出版する学術雑誌では、もちろん機関リポジトリ等からの電子的発信を許諾するものもあるが、著者が投稿した論文を自分のウェブページや機関リポジトリから発信することを禁止する契約を持つところもある。したがって、これらのポリシーをデータベース化することが必要となる。個別のルーチンとしては、セルフアーカイブが許諾されていれば問題ないが、許諾されていない場合でも、大学等がリポジトリに搭載するときには、公衆送信権の利用者として許諾を得ればよい。そのためには3つの方法がある。

1. 著者が許諾を求める：権利者が文句を言いにくい、通常著者はそこまで手間をかけない。
2. 図書館が著者の委任を受けて許諾を求める：妥協点。
3. 図書館が許諾を求める：図書館としては登録作業がやりやすいが、本筋ではない。

さらに、登録・利用促進をめぐるマーケティングの問題がある。オンラインで論文を提供することにより、引用数が数倍となるという結果もある（それを否定する議論もある）。自分の研究成果をできるだけ多くの読者に届けることができるのが機関リポジトリの長所であるというのが説得のポイントではあるが、著者である研究者は多忙で、登録の手間をかけるのがなかなか難しいのが現状である。一方、機関リポジトリの利用を促進するためには、本文が読める論文など、データ本体にアクセスできるコンテンツが充実していなければならない。したがって、運用面とシステム面の双方からコンテンツを簡単に登録できるように改善する必要がある。

登録を簡単にするという以外にも、システム面で改善すべき点がある。まず学内では、大学の社会責任が強く要請されるようになるにつれ、研究者の管理

の透明性を高めることが急務となり、研究者の事務負担等研究外業務負担が増大している。機関リポジトリと業績管理システムを統合することにより、一括して機関リポジトリへの登録と業績リスト作成が可能となるようなシステムを開発し、業務負担を軽減することが望ましい。

また本来機関リポジトリではすべての教育研究成果を制限なくすべての人に発信するのが原則だが、知的財産の観点から、またセキュリティの観点から認証システムが要請されている。さらに、オープンコースウェアなど教育用情報発信システムとの連携が計画されている。

学外に目を向けると、研究成果がより多くの人目に留まるように検索エンジンとの連携が行われはじめている。また社会貢献のための地域連携も視野に入ってきた。さらに、国際的に複数の機関リポジトリを連携運用することにより、データグリッドとして研究基盤を構築する方向性もあらわれてきている。特に、研究データ型リポジトリでは、このような連携が本質的に重要であり、今後の開発が必要である。

参考文献

オープンアクセスジャパン「オープンアクセス機関リポジトリ関連年表」

<http://www.openaccessjapan.com/resources/history.html>

Registry of Open Access Repositories (ROAR)

<http://archives.eprints.org/>

Clifford A. Lynch, Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. *ARL Bimonthly Report* 226, February 2003.

<http://www.arl.org/newsltr/226/ir.html>

Raym Crow, The Case for Institutional Repositories: A SPARC Position Paper.

<http://www.arl.org/sparc/IR/ir.html> 日本語訳「機関リポジトリ擁護論 :

SPARC 声明書」

http://www.tokiwa.ac.jp/~mtkuri/translations/case_for_ir_jptr.html

国立国会図書館「電子図書館プロジェクト」2002年(2005年改定)

<http://www.ndl.go.jp/jp/aboutus/elib-project.html>

SPARC, 2006 SPARC Program Plan.

<http://www.arl.org/sparc/about/pp2006.html>

倉田敬子「オープン・アクセスは何をもたらすのか？ 学術情報流通における図書館の役割」

<http://www.openaccessjapan.com/resources/pdf/kurata2005mslis1.pdf>

サウサンプトン大学(国立情報学研究所訳)「Budapest Open Access Initiative (BOAI)のためのセルフアーカイビング FAQ」

<http://www.nii.ac.jp/metadata/irp/safaq/>

Martin Feijen, ワークショップ「研究成果ショウケースとしての機関リポジトリ:オランダ“Cream of Science”を中心に」でのコメント。千葉大学、2006年5月22日。