

③学術機関リポジトリとリソースオーガナイザに関する研究

図書館情報メディア研究科 宇陀 則彦・歳森 敦

1. はじめに

筑波大学附属図書館の電子図書館システムは、平成17年度にリプレース作業を行い、平成18年3月に稼動を開始した。新システムでは、学術情報ポータルと学術機関リポジトリを柱とし、横断検索、情報資源リンク機能、My Libraryなどの機能を新たに提供し、サービス向上を図る。その一方で、附属図書館は国立情報学研究所の「最先端学術情報基盤（CSI）の構築推進委託事業（機関リポジトリ構築・連携）」の委託を受け、学位論文を中心として機関リポジトリへの登録を推進しているところである。

本研究は、まず機関リポジトリ導入に関する調査を本学の教員全員（ただし、同時期に附属図書館名で行った別の調査の対象者は除く）に対して質問紙調査という形で実施した。2005年11月に調査票（1322通）を配布し、同年12月末日までに回収できた470票を有効回答とした。調査項目は、電子ジャーナルをはじめとする電子資料をどのような形で入手しているか、研究者が手元に持っているデータにはどのようなものが多いか、機関リポジトリを実施したときに登録する意思があるかどうかなどである。

次にリソースオーガナイザに関する2つの予備実験を行った。ひとつはリンクシステムとして有名なSFXを導入し、リンクシステムの効果について考察した。もうひとつは既に公開されている機関リポジトリからメタデータを収集することで、機関リポジトリの連携可能性について考察した。

2. 学術機関リポジトリに関する意識調査

2.1 調査の概要

本調査は研究開発室と本学知的コミュニティ基盤研究センター（同センター松林麻実子講師、図書館情報専門学群 竹入ふみ）の共同で実施した。調査の対象として、平成17年度筑波大学職員録から研究科専攻毎に客員、連携教員を除く準研究員以上の研究者を抽出した。附属図書館が前後して実施した別の調査の対象者を除外して、残る1322名全員を調査対象とした。学内便で質問紙を送付し、同封した返信用封筒で学内便によって返送するよう依頼した。また、学内便を使わずWebページ経由でも回答できるようにした。2005年11月16日（水）に調査票を発送し、11月28日（月）を回答締切日としたが、最終的に12月16日（金）までに到着した回答を集計の対象とした。有効回答数は470票（内、Web経由72票）であり、回収率は35.6%であった。研究科別の配布数、有効回答数と回収率を表2.1に示す。また、年齢階層別、職位別の回答者数を表2.2、表2.3に示す。研究に必要な情報の入手状況に関する質問については、紙数の制約のため割愛する。

表2.1 研究科別配布数と回収率

研究科	配布数	有効回答数	回収率	全体に対する割合
人文社会科学	256	59	23.0%	12.6%
ビジネス科学	56	15	26.8%	3.2%
数理工学	126	48	38.1%	10.2%
システム情報工学	221	86	38.9%	18.3%
生命環境科学	116	49	42.2%	10.4%
人間総合科学	487	190	39.0%	40.4%
図書館情報メディア	60	22	36.7%	4.7%
不明（無回答）		1		0.2%
合計	1322	470	35.6%	100.0%

表2.2 年齢階層別回答者数

20歳代	30歳代	40歳代	50歳代	60歳代	無回答	合計
13	121	175	126	33	2	470
2.8%	25.7%	37.2%	26.8%	7.0%	0.4%	100.0%

表2.3 職位別回答者数

教授	助教授	講師	助手	準研究員	その他	無回答	合計
155	140	120	26	20	8	1	470
33.0%	29.8%	25.5%	5.5%	4.3%	1.7%	0.2%	100.0%

2.2 研究成果の公表

回答者は大部分が学術雑誌論文として成果を公表しており、一部の研究科で少数が紀要論文、図書を選択している（表2.4）。また、公表するメディアを選択する理由は、「業績として認められ」ることと「他の多くの研究者に広める」ことが重視されており、「研究成果の質を保証する」という見方もある（表2.5）。

自身の情報発信の態勢に関しては、Webページを持っている人は43.0%で持っていない人（57.0%）を下回っている。Webページでは「論文リスト・業績リスト」（83.6%）や「研究室紹介」（66.2%）、「略歴」（59.2%）、「研究プロジェクト情報」（45.3%）などが公開されているが、「論文の全文情報」は25.4%と比較的少数に留まっており、研究者個人のレベルでは、既発表論文の公開（セルフアーカイビング）が必ずしも進んでいないことを示している。

表2.4 成果公開の手段（択一）

	1	2	3	4	5	6	7	無回答	合計
人文社会科学	45	4	0	6	0	0	0	4	59
	76.3%	6.8%	0.0%	10.2%	0.0%	0.0%	0.0%	6.8%	100.0%
ビジネス科学	13	1	0	0	0	0	1	0	15
	86.7%	6.7%	0.0%	0.0%	0.0%	0.0%	6.7%	0.0%	100.0%
数理物質科学	44	0	1	0	0	0	0	3	48
	91.7%	0.0%	2.1%	0.0%	0.0%	0.0%	0.0%	6.3%	100.0%
システム情報工学	70	0	1	0	5	0	2	8	86
	81.4%	0.0%	1.2%	0.0%	5.8%	0.0%	2.3%	9.3%	100.0%
生命環境科学	45	1	0	1	0	0	0	2	49
	91.8%	2.0%	0.0%	2.0%	0.0%	0.0%	0.0%	4.1%	100.0%
人間総合科学	162	6	0	4	3	1	3	11	190
	85.3%	3.2%	0.0%	2.1%	1.6%	0.5%	1.6%	5.8%	100.0%
図書館情報メディア	19	0	0	1	0	1	1	0	22
	86.4%	0.0%	0.0%	4.5%	0.0%	4.5%	4.5%	0.0%	100.0%
不明 (無回答)	1	0	0	0	0	0	0	0	1
	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%
合計	399	12	2	12	8	2	7	28	470
	84.9%	2.6%	0.4%	2.6%	1.7%	0.4%	1.5%	6.0%	100.0%

凡例：1 学術誌、2 紀要、3 プレプリント、4 図書、5 学会発表、6 報告書、7 その他

表2.5 成果公開手段の選択理由（択一）

	度数	相対度数
研究成果を他の多くの研究者に広めるため	119	27.9%
学術雑誌に掲載された論文でなければ業績として認められないため	172	40.4%
研究成果の先取権を主張するため	9	2.1%
審査を通して研究成果の質を保証するため	55	12.9%
研究者としての義務であると考えている	35	8.2%
紙数などの形式面での制約がないため	8	1.9%
特に理由はないが習慣としている	13	3.1%
その他	15	3.5%
回答数	426	100.0%

2.3 機関リポジトリへの態度

本学に機関リポジトリが設置された場合に、登録したいとする回答は半数を僅かに超え（50.4%）、わからないとする人（41.4%）を上回った。登録したくないとする回答者は少数（8.2%）であった。登録したいとする理由（表2.6；複数回答）は、「研究成果をより多くの人に公開でき」、「研究・教育資源の共有化に有効だから」である。「オープンアクセスの理念に賛同しているから」は21.3%であるが、提示した選択肢中では最も選択が少なかった。掲載したい内容は、「学術雑誌論文」（95.5%）、「紀要論文」（38.2%）、「各種報告書」（29.5%）の順であった。わからない、登録したくないとした回答者は、「研究成果はその他の方法で発表しており、それで満足」と考えており、「登録作業が面倒」で「そもそも学術機関リポジトリというものが…わからない」と思っている。

2.4 まとめ

筑波大学の大半の研究者は、業績として認められ、かつ他の研究者に成果を広めるために学術雑誌に研究成果を公表している。研究者の半数弱はWebページで研究を中心とした自らの情報を発信しているが、論文や報告書の全文情報の公開にまでは至っていない。

機関リポジトリの設置に対して、半数は登録したいと回答しているものの、残る半数はわからないとしており登録に消極的である。消極的な人々は、機関リポジトリへの登録が、研究成果の雑誌への公表という方法以上に新たな価値を生み出すことに懐疑的であると言えよう。登録したい人々は、多くの人への新たな公開手段として捉えており、登録に関する態度にはこうした意識の違いが反映していると思われる。また、消極的な人々の一部は、学術機関リポジトリ自体を「わからない」としており、広報および実サービスの提供による浸透は、こうした状況を改善するきっかけになるだろう。

登録したいとする人々は、機関リポジトリへの登録によって、研究成果がより多くの人にアクセスされることを期待しており、機関リポジトリの成否は、登録によって確実にアクセスが増加するという実績の達成にかかっている。したがって、外部の検索サービスへの登録情報提供の仕組みの構築、他機関の機関リポジトリを含む包括的な情報検索の枠組みの提供等によって、機関リポジトリに登録された成果物の露出度を向上させることが最優先の課題である。また、これに加えて、登録作業の負担の軽減と機関リポジトリそのものの広報活動を行うことで、現時点で消極的な研究者の参加を誘発する枠組みを作ることも必要であろう。

表2.6 機関リポジトリに登録したいと思う理由（MA）

	度数	
研究成果をより多くの人に公開できるから	198	84.3%
可視性が上がり、研究成果の被引用数が高くなるから	83	35.3%
研究・教育資源の共有化に有効だから	123	52.3%
研究成果を永続的に保存できるから	71	30.2%
大学による統一的な情報発信体制の確立が必要であると考えているから	56	23.8%
大学の知名度や評価を上げることにつながると思うから	53	22.6%
オープンアクセスの理念に賛同しているから	50	21.3%
その他	2	0.9%
回答者数	235	

表2.7 登録したくないまたはわからないと思う理由（MA）

	人数	構成比
著作権の問題が心配だから	74	32.5%
登録作業が面倒そうだから	95	41.7%
現在、研究成果はその他の方法で発表しており、それで満足だから	111	48.7%
利用者による情報の悪用が心配だから	42	18.4%
何を登録すれば良いのか分からないから	26	11.4%
オープンアクセスの理念そのものに賛同できないから	5	2.2%
そもそも学術機関リポジトリというものが何なのか良く分からない	80	35.1%
その他	32	14.0%
回答者数	228	

3. リソースオーガナイザ実現に向けて

3.1 リソースオーガナイザとは

研究教育に有用な情報資源はインターネット上に数多く存在していると言われるが、利用者が知っている情報資源はごく一部であり、実際に研究教育に役立てている情報資源の数はさらに少ないだろう。大学図書館は利用者の情報要求に応えるべく電子ジャーナルやデータベースを提供しているが、そのようなサービスが行われていることを知らない利用者、知っていても使わない利用者、使い方がわからない利用者などがいる現状から判断して、電子図書館サービスが十分機能しているとはいえない。さらに根本的な問題として、大学図書館は利用者が求める情報を本当に提供しているのかという問題がある。情報を探すときは「まずはGoogle」という風潮をみると、大学図書館の情報提供のあり方にどこか問題があると言わざるを得ない。利用者がGoogleだけで十分な情報を得られるならそれはそれでよいかもしれないが、Googleもまだ全ての情報を提供するレベルには達していないので、研究教育を支援する大学図書館の意義は失われていない。

リソースオーガナイザは、①どういう情報資源あるかわからない、②どこに情報資源があるかわからない、③情報資源の使い方がわからない、④情報資源への最適なアクセスパスがわからないという問題を解決するために、「最小限のアクションで最適な情報資源へ導くこと」を目指すものである。すなわち、単に情報資源のリストを作成しておわりにするのではなく、孤立している情報資源同士を丹念につなぎ、発生する多くのパスの中から最適なパスを提供するところまでオーガナイズすることを意味する。しかも、パスの意味づけは情報要求に応じて異なるので、利用者ごとに用意するところまでもっていきたい。情報を探すときは「まずは大学図書館」になることが目標である。

3.2 リンキングシステムSFXの導入効果

リソースオーガナイザを設計する前段階として、今年度はリンキングシステムとして有名なSFXを導入し、リンキングシステムの効果について考察した。SFXはデータベースやサーチエンジンの検索結果として表示される書誌情報から、論文本体（フルテキスト）への誘導を行うことで、検索から原文取得までをより円滑につなぐことができる。導入にあたっては、平成17年10月から平成18年2月までは知的コミュニティ基盤研究センターの経費で、平成18年3月以降は附属図書館電子図書館システムの経費でサービスを継続する。平成17年度で契約した情報資源は以下のとおりである。

リンク元	リンク先
Web of Science	Science Direct
First Search	Blackwell Science Synergy
Science Direct	Springer Link Journals
	OPAC
	NII Webcat
	Google Scholar
	Amazon.com

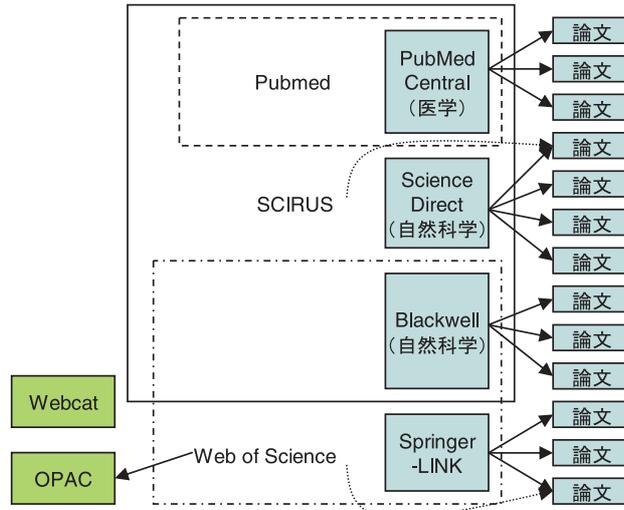
リンク元の情報資源を検索すると、検索結果の末尾に [SFX] と書いたボタンが示される。そのボタンをクリックすると、SFXメニューが別ウインドウで起動し、その論文本体を入手可能な複数の情報資源、すなわちリンク先を表示する。

SFXを導入することによるリンキング状態の変化を図3.1に示す。ここでは、SFXだけでなく、新電子図書館システムで導入する横断検索システムMetalibも含めて説明する。四角の包含関係は横断検索を表し、矢印は書誌情報から一次情報へのリンクまたは書誌情報から書誌情報へのリンクを表す。

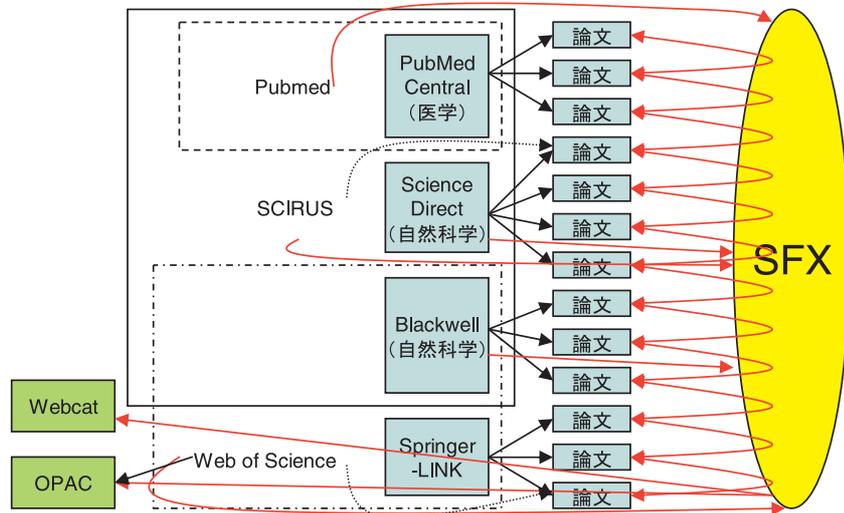
[SFX導入前]：電子ジャーナルやデータベースを個別に利用する状態

- Science Direct, Blackwellのサーバで論文を検索し、一次情報にアクセスする。
- SCIRUS, Web of Scienceから横断検索を行い、一次情報にアクセスする。(横断検索の対象およびリンクが部分的である。)

SFX導入前



SFX導入後
(平成17年10月)



Metalib導入後
(平成18年3月)

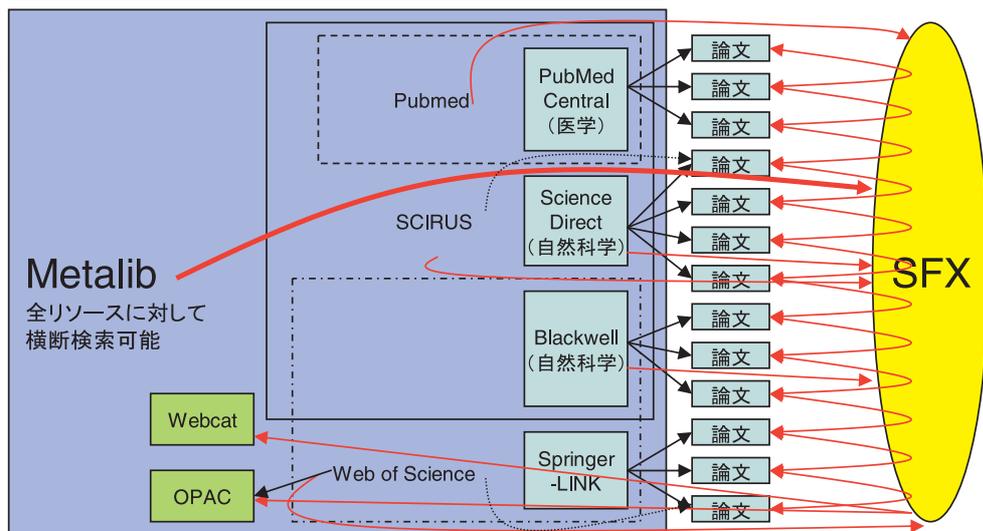


図3.1 リンキング状態の変化

- OPAC, Webcatを独立して使う。雑誌名までは検索できる。
- Web of ScienceからOPACのリンクがある。(ただし、雑誌名までのリンク)

[SFX導入後]：一つの電子ジャーナルやデータベースの検索結果を起点として、他の論文へリンクされている状態

- Web of ScienceやSCIRUSなど、データベースの検索結果から一次情報へのリンクが作成され、利用者を適切な情報源にナビゲートする。
- 論文の末尾にある参考文献リスト（書誌情報）から他の論文（一次情報）へリンクが作成され、検索し直さなくても、直接、一次情報にナビゲートする。
- 一次情報へのリンクだけでなく、OPACやWebcatやILLなど多様な情報資源へのリンクが生成され、しかも、それらをメニューから選択できる。

[Metalib導入後]：全てのリソースに対する横断検索の結果を起点として、他の論文へリンクされている状態

Web of ScienceやSCIRUSは限定された情報資源しか横断検索できないが、Metalibを導入することによって、ほとんど全ての情報資源を横断検索できる。また、検索結果をSFXに渡し、SFXを介して論文にアクセスできる。本実証実験によってリンクシステムが情報アクセスを円滑にすることが確かめられた。

3.3 機関リポジトリ連携におけるメタデータの利用

大学図書館が提供する情報資源は、電子ジャーナルやデータベースだけでなく、機関リポジトリに蓄積されたコンテンツも含まれる。今後、機関リポジトリが充実すれば、他の情報資源とリンクすることによってさらにコンテンツの価値はあがるだろう。逆にいえば、単一の機関リポジトリだけでは、大学の研究成果公開という意味での価値はあるものの、利用という意味での価値は低い。なぜなら電子ジャーナルやデータベースとは違い、主題や分野ごとにコレクションされているわけではないからである。しかし、もし登録されたコンテンツが関連する他の情報とリンクしていれば、付加価値が生じ、利用者がアクセスしたくなるような機関リポジトリになる可能性がある。

そこで本研究では、機関リポジトリ連携を想定し、既に公開されている機関リポジトリからメタデータを自動的に収集し、横断検索するシステムを構築することで、機関リポジトリ連携の可能性について考察する。なお、本研究は一部、科学研究費補助金「分散環境における異種メタデータの横断利用および学術機関リポジトリ運用に関する研究（基盤B）」と国立情報学研究所共同研究費「機関リポジトリ連携を前提とした情報資源リンクおよび知識探索パスの類型化」によった。

機関リポジトリ連携の第一段階として、まず複数機関リポジトリのメタデータを収集し、横断検索するための実験環境を構築した。実験環境には機関リポジトリ構築において広く使われているソフトウェアDSpaceを用い、メタデータハーベスティングプロトコルOAI-PMH（Open Archive Initiative: Protocol for Metadata Harvesting）を介してメタデータを収集した。DSpaceはマサチューセッツ工科大学(MIT)とヒューレットパッカーが共同開発したオープンソースのリポジトリ構築ソフトウェアである。DSpaceでは、標準で各種データ(PDF/HTML/JPEG/MP3等)の投稿機能、投稿の際に付与されたメタデータの検索に対応し、OAI-PMHに基づいたメタデータハーベスティングが可能である。また、SFXサーバが稼動していれば、メタデータを利用したOpenURLの表示が可能であり、送られてきたOpenURLは検索モジュールへ転送できる。リポジトリはコミュニティ単位で登録可能で、機関（組織）内の部局に対応しやすくなっている。また、コミュニティの中にさらに複数のコレクションを持つことができる。検索はコミュニティ全体を対象とした検索のほか、一つのコミュニティもしくは一つのコレクションを指定して検索できる。複数コミュニティあるいは複数コレクションは指定した検索はできない。

DSpaceには自動的にメタデータを収集する機能はないので、この部分は新規に作成した。メタデータの収集先は必要に応じて登録、削除が行えるようにし、一旦登録したならば、定期的にメタデータを収集する

ようになる。収集の周期も日にち単位で指定できる。収集したメタデータはDSpace内のDBMS PostgreSQLに格納するが、この際、メタデータスキーマを統一する必要があるため、各機関のメタデータスキーマをDSpaceのスキーマに変換する機能を実装した。収集および変換の結果(成功/失敗)は収集先ごとに電子メールで管理者に通知する。

なお、メタデータの収集先は原則OAI-PMHに則った機関とするが、その他のメタデータ、例えば書誌データのなものも取り込めるようにした。今回は実験的に日本図書館情報学会作成の文献目録データベースBIBLISと筑波大学知的コミュニティ基盤研究センター作成のULIS COREを対象とした。OAI-PMHでメタデータを収集できる機関は、400以上あり、Open Archive Initiativeのサイトにそのリストがある。

(<http://www.openarchives.org/Register/BrowseSites>)

現時点では、実験環境を構築したところまでで、収集したメタデータは2、3の機関にとどまっている(図3.2参照)。今後、400機関全てのメタデータを収集し、メタデータの関連度などを分析しながら、機関リポジトリ連携の効果について検証したい。



図3.2 Dspaceのメイン画面

3.4 今後の予定

1章冒頭で述べたとおり、3月稼動の新電子図書館システムでは、学術情報ポータルと学術機関リポジトリを柱とし、横断検索、情報資源リンク機能、My Libraryなどの機能を新たに提供する。横断検索にはMetalib、情報資源リンクにはSFX、そして学術機関リポジトリにはDSpaceがそれぞれ導入される。横断検索と情報資源リンクに関しては、平成17年度で対象とした情報資源から大幅に数を増やし、筑波大学で契約しているほとんど全ての情報資源を対象にする予定である。来年度は横断検索と情報資源リンクのサービス状況を観察しながら不足している機能や改善すべき機能を洗い出し、さらに高度なサービス機能を実現したいと考えている。また、機関リポジトリに関しては、電子図書館システムで導入するDSpaceに今年度研究開発室で作成した自動登録機能を追加し、機関リポジトリ連携を実現したい。

実装の話としては、昨今はやりのWeb関連技術を積極的に取り入れる予定である。

- OAI-PMH, DCMES, Application Profile
- OpenURL, Link Resolver (Open resolver etc.)
- SRU/SRW, Web Service (SOAP, WSDL, UDDI)
- Semantic Web (OWL, SKOS)

これらの技術を組み合わせ、うまく汎用性の高いモジュールが開発できれば、他大学の図書館システムや国立情報学研究所システムとの相互接続が容易になるだろう。

4. おわりに

ニーズがあるからシステムができるのか、システムができたからニーズが生まれるのか。機関リポジトリは明らかに後者を狙うしかない。「機関リポジトリが何なのかが分からない」という声に対して、説明や広報だけでは不十分である。百聞は一見にしかず、機関リポジトリの姿をまず見せることである。ただし、下手に見せると逆効果にしかならないので、どう見せるか十分な検討が必要である。機関リポジトリを単独で見せることはせず、利用者は何を望むのかという視点にたって、図書館サービス全体の中で位置づけることが重要である。個々のコンテンツ、個々の機能に限定して論じても意味がない。図書館はコンテンツの集合体であるコレクション、機能の集合体であるシステムで勝負である。リソースオーガナイザはその具体的なかたちである。