



Title	機関リポジトリへのアクセス経路
Author(s)	紙谷, 五月; 野中, 雄司; 杉田, 茂樹
Description	特集 ファインダビリティ向上
Citation	情報の科学と技術, 58(12), 610-614
Issue Date	2008-12-01
Doc URL	https://hdl.handle.net/2115/34975
Type	journal article
File Information	AIRway.pdf



機関リポジトリは独自の集客力を持たず、外部の情報サービスシステムに集客を依存する。従って、より多く利用されるためには、主たる閲覧利用者層として想定される研究者の日常的な情報探索行動の中で、収載文献が自然に発見されアクセスされるような経路の形成と誘導が必要となる。これを目的として実装され、リンク・リゾルバのための追加情報源として機能する AIRway システムを概観し、現在の被利用状況について述べ、今後の展望についてまとめる。

An institutional repository does not have its own power to attract large users, but commends user navigations to external service providers. Academics are assumed to be the primary users of institutional repository. Therefore, to gain their access, it is important to make the repository contents discoverable and accessible within their day-to-day research behavior. We have implemented “AIRway” system which is designed to be yet another knowledgebase for link-resolvers. This paper describes its functions and the usage.

キーワード：機関リポジトリ, リンク・リゾルバ, OAI-PMH, OpenURL, AIRway

Keywords: institutional repository, link-resolver, OAI-PMH, OpenURL, AIRway

1. はじめに

1990年代後半のインターネットの発達と電子ジャーナルの成立により、研究論文をはじめとした学術文献の流通形態はインターネットを介した電子流通に移行した。さらに現在、学術出版社によるジャーナルの出版流通に加え、大学や研究機関のウェブサイトからも所属研究者の研究論文を公開する動きが活発になってきている。こうしたウェブサイトは「機関リポジトリ」と呼ばれ、世界で約 600、我が国でも 90 近い大学や研究機関がこれを設置している。

機関リポジトリの主要な目的のひとつは、所属研究者の著作文献を広くインターネット公開することによって、より広汎な読者を獲得することにある。したがって、より多くの文献データベースやサーチエンジンから機関リポジトリ上の文献が検索・閲覧可能となることが望ましい。本稿では、機関リポジトリの集客手段の現況について概括した上で、リンク・リゾルバを通じた利用者誘導方式の実装例である AIRway システムについて述べ、その現況を示す。

2. 機関リポジトリの集客

機関リポジトリは独自の集客力を持たない。arXiv.org¹⁾や CERN Document Server²⁾, RePEc³⁾に見られるような分野的特徴に基づく求心力を備えているわけではなく、また、Google Scholar⁴⁾やトムソン社 Web of Science⁵⁾, エルゼビア社 Scopus⁶⁾などの全方位型

の学術文献データベースのような網羅性を有しているわけでもない。機関リポジトリには、さまざまな学術分野の、さまざまタイプの文献、講演スライド、データセットなどが、初学者向けのテキストから先端的な研究論文まで幅広く、しかし単一機関の所属研究者の著作に限って収録されている。ある学術的関心を持つ閲覧利用者が、数百とある世界中の機関リポジトリに順次アクセスし、文献探索を重ねるというのは合理的なやりかたではない。

特徴がなく、網羅性もないという点はしかし機関リポジトリの欠陥ではない。設置主体である大学や研究機関自身は情報サービス事業者ではなく、従って、コレクションとしての機関リポジトリという箱そのものに脚光が当たる必要はない。むしろそこで公開されるひとつひとつの文献が、各々の価値にふさわしい、望みうる最大の読者を得ることが機関リポジトリの意義の中心である。

機関リポジトリは、外部情報サービスに集客を依存する。OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)⁷⁾によるメタデータ・ハーベスティングや、ウェブクロールリングを受けることによって、OAI サービス・プロバイダやインターネット検索エンジンが、機関リポジトリに登録された資料の検索と利用のためのインターフェースとなる。

さてしかし、ミシガン大学による OAIster⁸⁾をはじめとした OAI サービス・プロバイダ、Google をはじめとしたインターネット検索エンジンらは、研究者を主たる読者として想定した場合、有効な集客手段として機能しているだろうか。

——機関リポジトリに論文を登録すると、Google などからアクセスできるようになります。

——私たちは文献を探すときは普通、Web of Science とか、PubMed⁹⁾を使うのですが、それらからはアクセスできないのでしょうか？

平成 17 年 8 月 22 日に北海道大学大学院医学研究科で催した、同大学機関リポジトリの学内説明会での質疑である¹⁰⁾。既存の OAI サービス・プロバイダの認知度は未だ低く、一般的なインターネット検索エンジンは学術文献の探索という用途には通常用いられない。とすると、研究者の自然な情報探索行動の中に、機関リポジトリ上の文献発見を位置づけていくためにはどのような方法があるだろうか。

平成 18 年 4 月、エルゼビア社は、ウェブ資源を含めた学術文献検索サービスである Scirus¹¹⁾の索引に、千葉大学の機関リポジトリである CURATOR を加えた¹²⁾。平成 19 年 2 月には北海道大学、早稲田大学、同 5 月には金沢大学、九州大学、熊本大学と Scirus の索引対象は多くの機関リポジトリに拡大している。

また、平成 20 年 10 月からは国立情報学研究所の CiNii が、OAI サービス・プロバイダとしての機能を新たに備え、学会誌等をはじめとした従来のコンテンツに加え、国内の大学、研究機関の機関リポジトリに収載された学術論文が検索可能となった¹³⁾。

これらは、まさに、研究者の自然な情報探索行動の中に、機関リポジトリ上の文献の発見を位置づけていく枠組みであるといえる。本稿では、さらなる別の試みとして、機関リポジトリ収載文献の情報を集約し、リンク・リゾルバの追加ナレッジベースとして機能する AIRway システムについて記す。AIRway は国立情報学研究所次世代学術コンテンツ基

盤共同構築事業学術機関リポジトリ構築連携支援事業¹⁴⁾の委託課題領域 2 として、筆者の所属する北海道大学が中心となってすすめる「機関リポジトリ上の情報資源の発見及びアクセス性の向上のための調査研究開発」¹⁵⁾の一環として実装されたものである。

3. AIRway とは

文献データベースの検索結果から電子ジャーナル上のフルテキストへの利用者誘導のための汎用的な手段として、リンク・リゾルバが用いられている。

リンク・リゾルバは、CrossRef¹⁶⁾等を集積された電子ジャーナルの書誌所在情報と、利用者所属機関の購読契約情報をもとに、アクセス可能なフルテキストの所在を利用者に示す中間システムである。AIRway は、各機関リポジトリに収録されたオープン・アクセス文献の情報を集積し、リンク・リゾルバの追加情報源として機能することを目指すバックエンドシステムである¹⁷⁾。

これによって、電子ジャーナル非購読契約機関にあっても、代替として、機関リポジトリにセルフアーカイブされたオープン・アクセス版のフルテキストに容易にアクセスできることとなる（図 1）。

AIRway は要素技術として OAI-PMH と OpenURL¹⁸⁾を用いる。OAI-PMH によって各機関リポジトリの収録文献情報を収集し、リンク・リゾルバからの OpenURL0.1 及び同 1.0 による問い合わせに対して、該当文献の書誌所在情報を返戻する（図 2）。

OpenURL リクエストに対する照合結果の返戻は、独自に定義した XML スキーマ ir.xsd¹⁹⁾によって行う。ir.xsd はジャーナル掲載文献記述のためのスキーマ ofi/fmt:xml:xsd:journal²⁰⁾に、機関リポジトリ及び本システム特有の必要情報を追加して拡張したものである（図 3、表 1）。

特徴的な拡張部分について以下に示す。

- (1) url : 該当文献のフルテキストを入手可能な URL
- (2) repository : 該当文献を収録する機関リポジトリ名称
- (3) resource_version : 該当文献のバージョン（刊行版であるか、著者原稿であるか）
- (4) identifier : 該当文献の識別子

該当文献の書誌情報は、metadata 要素以下で示される。metadata 要素は、現在 jnl:journal によって記述されるが、将来 bk:book, dis:dissertation への拡張の余地を残している。

4. AIRway の現況

本稿執筆時現在、リンク・リゾルバとしては、ExLibris 社 SFX²¹⁾の国内の実装及び OCLC WorldCat Link Manager（旧 1Cate）²²⁾が、AIRway に対応している。対象としている機関リポジトリは国内 13 大学及び英国 1 大学の合計 14 大学であり、対応論文数は 94,146 編である。

平成 18 年 11 月 1 日から平成 20 年 8 月 31 日までの、AIRway サーバへのアクセスログ

を解析したところ、OpenURL リクエストのうち、対象機関リポジトリに論文が収録されていて所在を解決できたのは 1.45%であった（表 2）。ただし、対象リポジトリ数、対応論文数がまだ僅少であることを考え併せるとこれは破格のヒット率であり、業界関係者等からの試験的アクセスが若干程度あったことを見込まねばならないだろう。

所在解決できたケースの内、OpenURL のクエリ中に含まれる rfr_id (参照元を示す識別記述子) から判明したリンクソースを参考までに図 4 に示す。

AIRway で所在解決されたもの（以下、「リゾルブ数」という。）のうち、実際にどの程度の利用者が機関リポジトリにアクセスした（以下、「実効リゾルブ数」という。）だろうか。

AIRway システムのログからは実効リゾルブ数は把握できない。そのため AIRway 対象機関リポジトリのひとつである北海道大学学術成果コレクション²³⁾（以下、「HUSCAP」という。）のログ（次章に詳述する）から得られた AIRway 経由のアクセス 333 件（アクセス元ホスト名未解決含む）との比較対照に基づき、全体の実効リゾルブ数を推定する。

AIRway 収録データ中に占める HUSCAP の論文数（21,762 編）比率から、実効リゾルブ数比率は約 7%と推定される（表 3）。

ほとんどのケースでは購読電子ジャーナルを閲覧したものと推定されるが、AIRway は電子ジャーナルへのアクセス権限のない潜在的読者を機関リポジトリに導くことにより、対応リンク・リゾルバの導入機関に対する当該文献の可視性を 7%向上させたと言える。

5. 機関リポジトリ利用に占める AIRway の位置

次に、機関リポジトリへのアクセス経路についての傾向およびその中に占める AIRway の位置について示す。

以下では、HUSCAP のアクセスログ（平成 18 年 11 月 1 日～平成 20 年 8 月 31 日）を用い、同リポジトリに収録された、AIRway の主ターゲットとなる文献群についてアクセス経路を分析した。

前処理として、論文書誌表示画面へのアクセスを対象として、次の手順で分析対象データの抽出を行った。

- (1) ウェブクローラによる系統的アクセス²⁴⁾等をはじめとした集計に適さないログデータを排除するため、COUNTER (Counting Online Usage of NeTworked Electronic Resources)²⁵⁾及び佐藤の方法²⁶⁾によりフィルタリング処理を行い、
- (2) 次に、リンクソースの大半は英文の学術雑誌論文を中心とした文献データベースであることから、分析対象を英文学術雑誌論文²⁷⁾に限定し、
- (3) 最後に、来着経路の分析のため、アクセスログ中の referer が確認できないケース及び HUSCAP 内からの簡略書誌画面へのアクセスを除外した。

集計結果を表 4 に示す。

抽出されたアクセスログを元に、教育機関およびそれ以外からのアクセス経路を比較する²⁸⁾。具体的には教育機関、教育機関以外を比較し、研究者からのアクセス経路の特徴、及び研究者と研究者以外からのアクセス経路に差異があるかどうかを確認した（図 5, 6）。

図中に示したアクセス経路の名称の意味は以下のとおりである。

- (1) サーチエンジン：Google 等サーチエンジンの検索結果からのアクセス。
- (2) 学術系サイト：Scirus や Scientific Commons ²⁹⁾などの学術情報ポータルサイト等からのアクセス。うち、"AIRway"が AIRway からのアクセスである。
- (3) その他：ブログやウィキペディア，研究者個人のホームページ等からのアクセス。

全体的な傾向としては，教育機関とそれ以外を通じ，サーチエンジンを経由してのアクセスが多く，AIRway を通じたアクセスは教育機関からで 6%，教育機関以外からで 2%であった。

先述の通り，AIRway 対応リゾルバはまだ 2 種だけである。SFX の導入機関は世界中で 1,500（同ウェブサイトによる。以下同じ。うち国内の実装のみ AIRway 対応），360 Link を擁する Serials Solution のサービス導入機関は世界で約 2,000 とされる。今後，各リンク・リゾルバの AIRway 対応がすすむにつれ，この，機関リポジトリへのアクセス経路に占める AIRway の割合は増加していくことが見込まれる。

6. おわりに

本稿は，対応リンク・リゾルバの種類数からも，対象機関リポジトリ数からも，まだ生まれたばかりとも言える現在の AIRway の，初めてのログ分析結果である。

機関リポジトリの収載文献数は年々着実に増え続けている。各種文献アーカイブとの比較を図 7 に示す。円の面積が収載文献数の規模を表す。世界の機関リポジトリ収載文献数については英国サウサンプトン大学による調査 ³⁰⁾を，国内の機関リポジトリ収載文献数については国立情報学研究所による調査 ³¹⁾を，その他のプラットフォームについては各々のウェブサイトに示された数値を用いて比較している。

機関リポジトリの収載文献数は，arXiv.org や J-STAGE ³²⁾を超え，CiNii に匹敵する規模となっている。これはエルゼビア社の電子ジャーナル (ScienceDirect) ³³⁾の 3 分の 1 に相当する数字であり，オープン・アクセス文献は一定程度以上の規模と価値を持ちつつあると見ることができる。

対応するリンク・リゾルバの種類数と機関リポジトリ収録論文数が増加することによって，AIRway によるリゾルブ事例は増すものと考えられる。オープン・アクセス文献へのナビゲーションを通じた学術情報流通環境の向上に向け，本プロジェクトへの業界関係諸賢のさらなる理解と協力を望みたい。

しかしまた一方で，前段に挙げたようなさまざまな電子コンテンツの適切な組織化と統合的利用環境の実現のためには，リンク・リゾルバを含む現存の技術は過渡的なものとみただろうがよいかもしれない。研究者の情報探索行動に見合ったさらなるシステム改善，新たな枠組みの創出への努力が常に必要であると本プロジェクトは考える。

参考文献・注

各ウェブサイトの確認期日は全て[accessed 2008-09-24]である。

- 1) arXiv.org
<http://arxiv.org/>
- 2) CERN Document Server
<http://cdsweb.cern.ch/>
- 3) RePEc: Research Papers in Economics
<http://repec.org/>
- 4) Google Scholar
<http://scholar.google.com/>
- 5) Web of Science
<http://www.thomsonscientific.jp/products/wos/>
- 6) Scopus
<http://www.scopus.com/>
- 7) The Open Archives Initiative Protocol for Metadata Harvesting
<http://www.openarchives.org/OAI/openarchivesprotocol.html>
- 8) OAIster
<http://www.oaister.org/>
- 9) PubMed
<http://www.ncbi.nlm.nih.gov/pubmed/>
- 10) 杉田茂樹. “リンクリゾルバを通じた機関資源へのアクセス”. 国立情報学研究所平成18年度 CSI 委託事業報告交流会予稿集. 東京, 平成19年7月3日, 国立情報学研究所, 2007, p. 75-81.
<http://www.nii.ac.jp/irp/event/2007/debrief/pdf/4-1hokudai.pdf>
- 11) Scirus
<http://www.scirus.com/>
- 12) Scirus Builds on Global Success of Repository Search Service with the First Asia-Pacific Partner, Chiba University
http://www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_00446
- 13) CiNii と機関リポジトリの連携について
http://www.nii.ac.jp/irp/2008/09/cinii_1.html
- 14) 学術機関リポジトリ構築連携支援事業
<http://www.nii.ac.jp/irp/>
- 15) AIRway Project
<http://airway.lib.hokudai.ac.jp/>
- 16) crossref.org
<http://www.crossref.org/>
- 17) Sugita, S.; Horikoshi, K.; Suzuki, M.; Kataoka, S.; Hellman, E.S.; Suzuki, K. Linking Service to Open Access Repositories. D-Lib Magazine. 2007, volume 13,

Number 3/4

<http://www.dlib.org/dlib/march07/sugita/03sugita.html>

- 18) ANSI/NISO Z39.88-2004 The OpenURL Framework for Context-Sensitive Services

<http://www.niso.org/standards/z39-88-2004/>

後出の `rfr_id` についても同仕様に記載あり。

- 19) ir.xsd

<http://eprints.lib.hokudai.ac.jp/ir.xsd>

- 20) XML Schema defining the XML Metadata Format to represent serially published documents, and its component parts "issue" and "article"

<http://www.openurl.info/registry/docs/xsd/info:ofi/fmt:xml:xsd:journal>

- 21) SFX

<http://www.exlibrisgroup.com/category/SFXOverview>

- 22) WorldCat Link Manager

<http://www.oclc.org/linkmanager/>

- 23) 北海道大学学術成果コレクション

<http://eprints.lib.hokudai.ac.jp/>

- 24) 既知のロボットに加え，機関リポジトリ特有の系統的アクセス（機関リポジトリ調査機関等によるものなど）もここにふくめた。

- 25) COUNTER (Counting Online Usage of NeTworked Electronic Resources).

<http://www.projectcounter.org/>

- 26) 佐藤義則. 機関リポジトリの利用統計のゆくえ. カレントアウェアネス.2008 , no296, p.14-16.

- 27) HUSCAP に登録されている英文学術雑誌論文は，2,361 文献である。

- 28) 教育機関はトップレベルドメイン，セカンドレベルドメインから教育機関と判別可能なドメインのみを教育機関とみなした。国によっては教育機関を表すドメイン名にまったく規則性がないことも多く，（たとえば日本では主に `ac.*` のような規則性がある）そのような教育機関は抽出できていない。そのため，世界中の教育機関を抽出できているわけではないが，教育機関と判別できるドメインのみを抽出し，教育機関すなわち研究者の特徴を把握することを目的としたためこのような手法をとった。

- 29) Scientific Commons

<http://www.scientificcommons.org/>

- 30) Registry of Open Access Repositories (ROAR)

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

- 31) IRDB コンテンツ分析システム

<http://irdb.nii.ac.jp/analysis/index.php>

- 32) J-STAGE

<http://www.jstage.jst.go.jp/>

- 33) ScienceDirect

<http://www.sciencedirect.com/>

図1: リンクリゾルバを通じたフルテキストへの想定アクセス経路

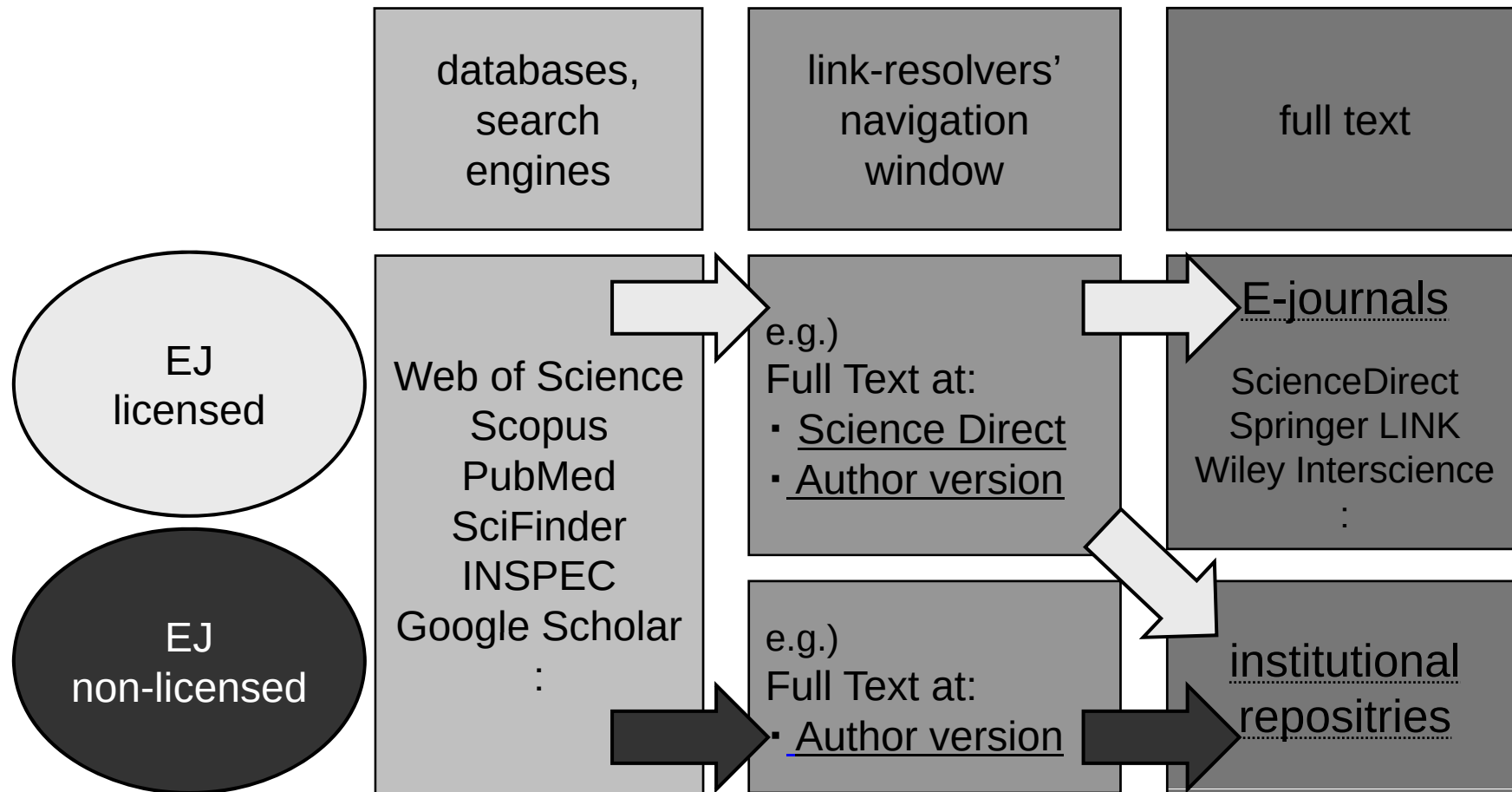


図2: AIRwayの動作イメージ

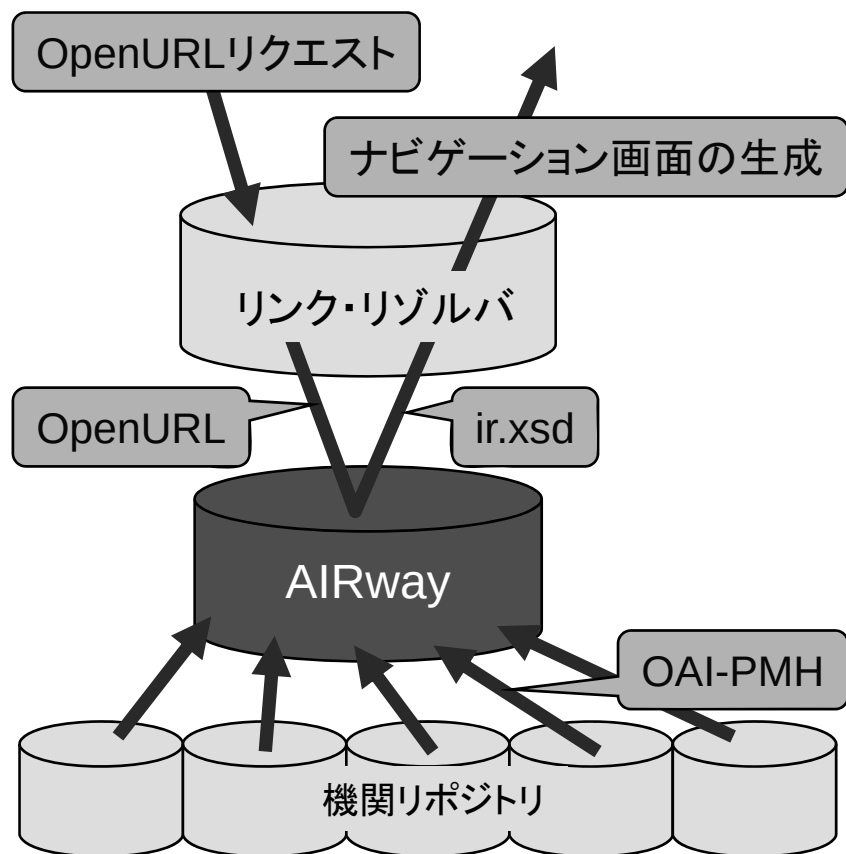
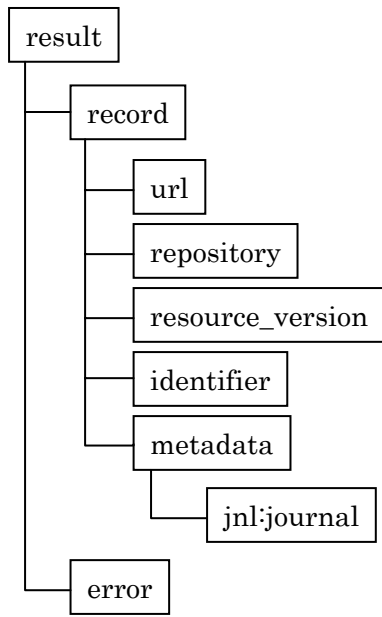


図 3 : ir.xsd



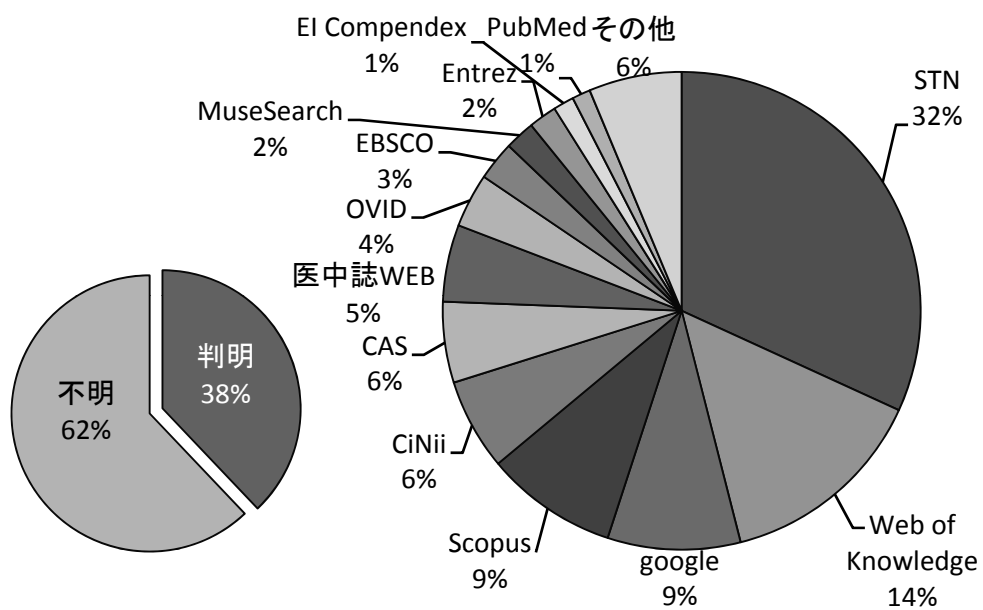


図4: 主なリンクソース

図5 教育機関からのアクセス経路

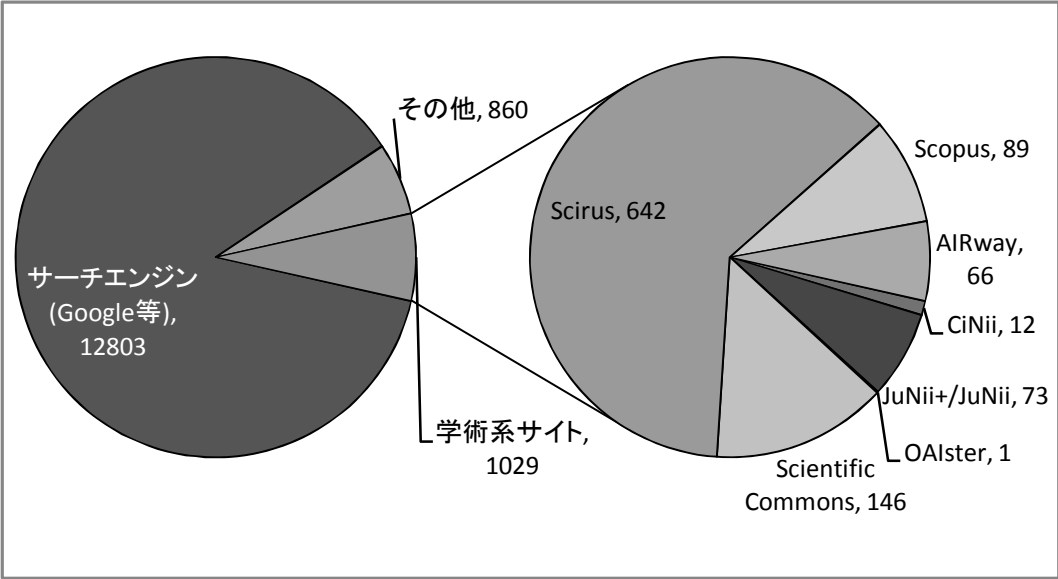
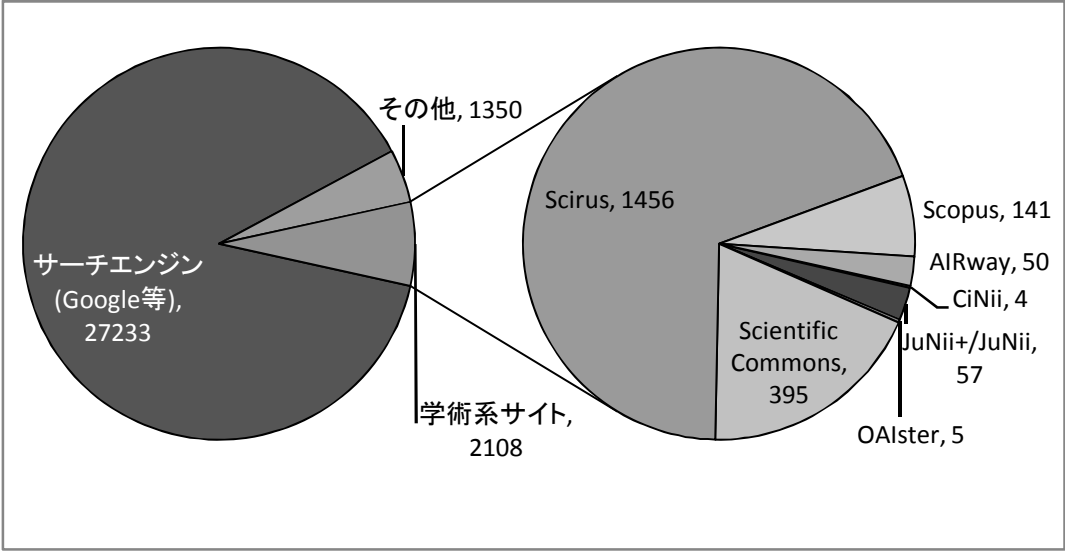
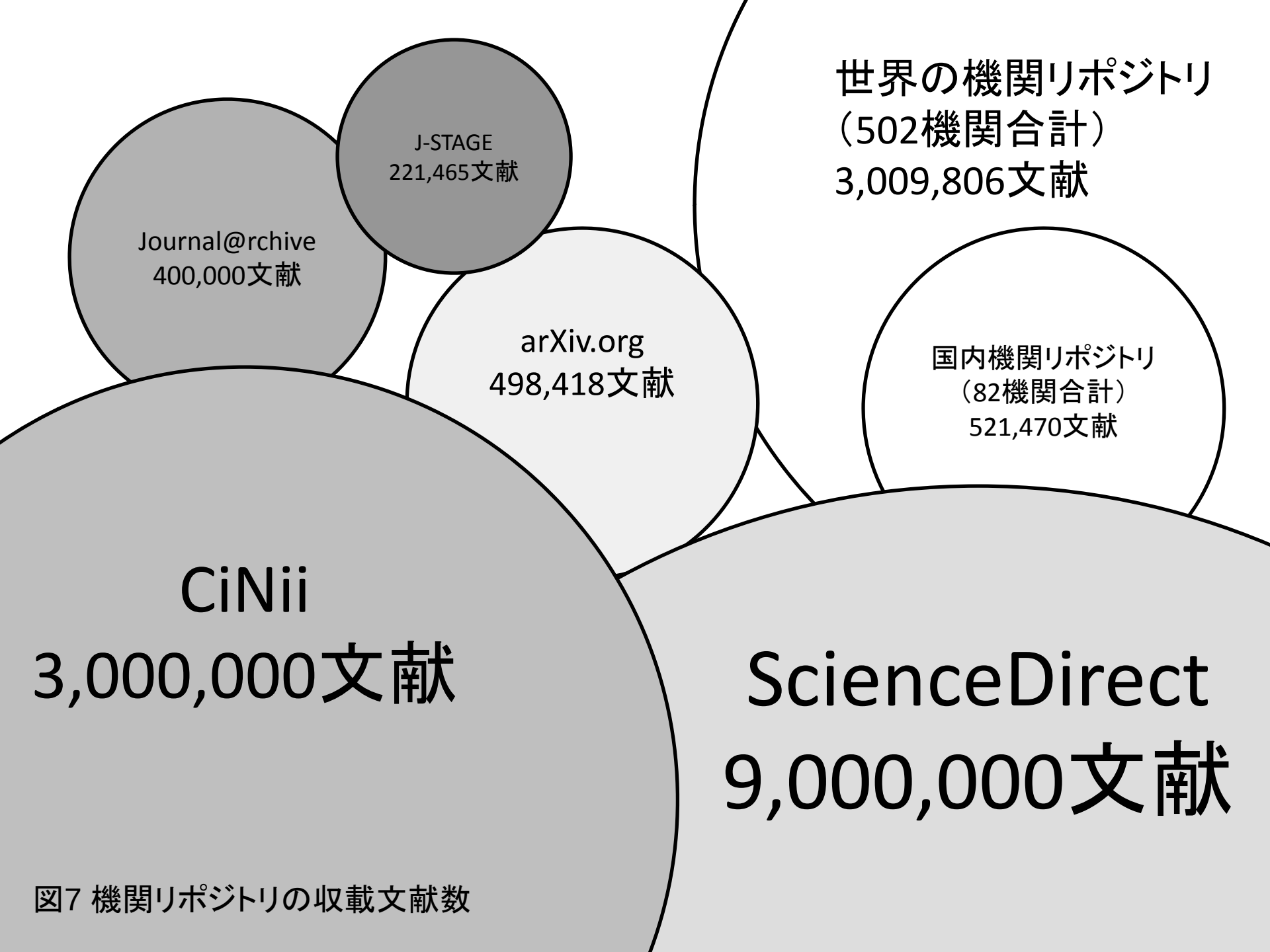


図6 教育機関以外からのアクセス経路





世界の機関リポジトリ
(502機関合計)
3,009,806文献

国内機関リポジトリ
(82機関合計)
521,470文献

arXiv.org
498,418文献

Journal@rchive
400,000文献

J-STAGE
221,465文献

CiNii

3,000,000文献

ScienceDirect
9,000,000文献

図7 機関リポジトリの収載文献数

表 1 . AIRway の返戻する XML の例

```
<?xml version = "1.0" encoding = "UTF-8"?>
<result xmlns="http://eprints.lib.hokudai.ac.jp/ir"
xmlns:jnl="info:ofi/fmt:xml:xsd:journal"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://eprints.lib.hokudai.ac.jp/ir
http://eprints.lib.hokudai.ac.jp/ir.xsd" version="1.0"
status="resolved">
<record hitno="1">
  <url>http://hdl.handle.net/2115/559</url>
  <repository>Hokkaido University Collection of Scholarly
and Academic Papers</repository>
  <resource_version>author</resource_version>
  <identifier>info:doi/10.1002/jez.a.205</identifier>
  <metadata>
    <jnl:journal>
      <jnl:authors>
        <jnl:au>Kawamoto, Shishin</jnl:au>
        <jnl:au>Yoshida-Noro, Chikako</jnl:au>
        <jnl:au>Tochinai, Shin</jnl:au>
      </jnl:authors>
      <jnl:atitle>Bipolar Head Regeneration Induced by
Artificial Amputationin Enchytraeus japonensis (Annelida,
Oligochaeta)</jnl:atitle>
      <jnl:jtitle>Journal of Experimental Zoology Part A:
Comparative Experimental Biology</jnl:jtitle>
      <jnl:date>2005-07-12</jnl:date>
      <jnl:volume>303A</jnl:volume>
      <jnl:issue>8</jnl:issue>
      <jnl:spage>615</jnl:spage>
      <jnl:epage>627</jnl:epage>
      <jnl:issn>15488969</jnl:issn>
    </jnl:journal>
  </metadata>
</record>
</result>
```

論文の所在解決を目的としたOpenURLリクエスト	1,484,923
所在が解決できたOpenURLリクエスト	21,599

表2. AIRwayの所在解決回数

	HUSCAP	AIRway全体
収録論文数	21,762	94,146
リゾルブ数	4,993 (推定)	21,599
実効リゾルブ数	333	1,441 (推定)

表3. AIRwayの実効リゾルブ数

論文書誌画面への総アクセス件数	4,679,456
(1)フィルタリング後のアクセス件数	1,020,920
(2)英文学術雑誌論文へのアクセス件数	121,287
(3)判別可能な外部サイト経由アクセス件数	45,383

表4. HUSCAP論文書誌表示画面へのアクセス数