

デジタルリポジトリオーバーレイによって実現する Digital Mathematics Library

行木孝夫¹, 黒田 拓²

E-mail: {nami, hiraku}@math.sci.hokudai.ac.jp

1. 北海道大学大学院理学研究院数学部門 (数学連携研究センター兼務)
 2. 北海道大学大学院理学院数学専攻
- 〒060-0810 北海道札幌市北区北10条西8丁目

概要

日本から発行されている数学系ジャーナル類はデジタルリポジトリによって電子化されたタイトルが数多く存在する。これらを統合し、世界的な Digital Mathematics Library の枠組みへ寄与する試みを紹介する。

キーワード

OAI-PMH, Digital Mathematics Library, digital repository, repository overlay.

Digital Mathematics Library over digital repositories

Takao Namiki¹ and Hiraku Kuroda²

E-mail: {nami, hiraku}@math.sci.hokudai.ac.jp

1. Department of Mathematics, Hokkaido University
 2. Graduate School of Science, Hokkaido University
- Kita 10 Nishi 8, Kita-ku Sapporo 060-0810, Japan

Abstract

There are huge number of mathematical journal titles in Japan and many of electronic edition of these titles are loaded on digital repositories. In this paper the authors introduce an experimental subject portal which is joined with framework of World Digital Mathematics Library.

Keywords

OAI-PMH, Digital Mathematics Library, digital repository, repository overlay.

1 背景

日本における国際誌としての数学系ジャーナルは数学教室から発行されるタイトルが数百にのぼり [2]、コミュニティとしての出版形態が主流である。総論文数は約 70,000 に達する。タイトルの一部を付録 A に示した。これらは全て Math. Reviews (MR) に収録、データベース化され、世界の研究者から先行研究を調査する際の対象となっている。我々は、Project Euclid および機関リポジトリから取得できるメタデータを利用して数学系ジャーナルのポータルを形成した。結果、各国で構築が進む Digital Mathematics Library の日本パート (DML-JP) を実現しつつある。数

学コミュニティにおける日本の責任を果たす成果として、国際的にも評価されつつある。 α 版の URL は <http://dmljp.math.sci.hokudai.ac.jp/>、 β 版の URL は <http://sparc1.math.sci.hokudai.ac.jp/dmljp> である。日本数学会のドメインで同じサービスを提供する予定もある。

1.1 収録対象誌

SPARC Japan パートナーである Project Euclid から 10 タイトル、CSI 委託事業の対象である機関リポジトリから 17 タイトル、計 27 タイトルを収録した。タイトルのリストを下記に示す。把握した論文数はおよそ 30,000 である。

Bull. Tokyo Gakugei University Sec. I
Bulletin of College of Science the University Ryukyu
Hiroshima Math. J.
Hokkaido Mathematical Journal
J. Math. Soc. Japan
Japan J. Indust. Appl. Math.
Journal of Mathematical Sciences, The University of Tokyo
Journal of the Faculty of Education, Kagoshima University
Journal of the Faculty of Science Shinshu University
Journal of the Faculty of Science, Kagoshima University
Journal of the Faculty of Science, the University of Tokyo Sect 1 A
Journal of the Faculty of Science, Yamagata University
Nagoya Math. J.
Kodai Math. J.
Nat. Sci. J. Fac. Educ. Hum. Sci. Yokohama National University Sec. I
Natur. Sci. Report. Ochanomizu. Univ.
Nihonkai Mathematical Journal
Osaka J. Math.
Proc. Japan Acad. Ser. A Math. Sci.
Publ. Res. Inst. Math. Sci.
Reports of the Faculty of Science and Engineering, Saga University. Mathematics
RIMS Kokyuroku
Ryukyu Mathematical Journal
Sci. Rep. Yokohama National University Sec. I
The science reports of the Kanazawa University
Tohoku Math. J.
Tokyo J. of Math.
Tsukuba Journal of Mathematics

2 実装

メタデータハーベスティングとメタデータ変換、Math. Reviews とのマッチング、サービスを提供するインターフェースとに分け、各ステップを順次実装した。対象とするリポジトリは Cornell 大学図書館の運営する数学及び統計学分野の電子ジャーナルポータル Project Euclid を始め、機関リポジトリ、arxiv.org とした。機関リポジトリの選択に当たってはブラウズページを目視で閲覧し、数学系ジャーナルが搭載されているかどうかを確認している。

2.1 メタデータハーベスティング

メタデータハーベスティングに際しては、プラットフォームに応じて oai_dc および junii2 形式のメタデータ型を選択した。これは、OAI-PMH における metadataPrefix として設定する。

Project Euclid[4] は oai_dc のみ提供している。各論文の書誌情報を次のように dc:identifier タグ内に記述しており、ここからジャーナルタイトル、巻号、発行年、ページ範囲を抽出する作業は Project Euclid 側で形式が変わるごとに対応しなければならない。Math. Reviews によって論文へ付与される主題分類は dc:subject タグ内に記述している。ジャーナルタイトルは Math. Review に準拠しており、マッチングの際に処理が容易である。

```
<dc:identifier>J. Math. Soc. Japan 25, no. 1 (1973), 1-6</dc:identifier>
```

国内の機関リポジトリが提供する junii2 フォーマットにおいて書誌情報は必要な粒度でタグ付けされており、Project Euclid のような困難はない。一方で、ジャーナルタイトルが日本語のみ提供される例が見られるなど、サブジェクトベースのリポジトリとは異なる困難が存在する。

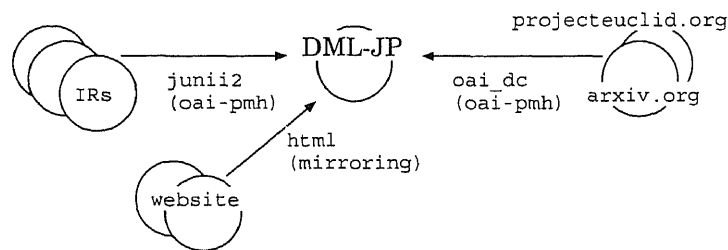


図 1: メタデータハーベスティング部分の概念図

メタデータに限らず、オープンアクセスなコンテンツに関してはフルテキストをミラーリング、格納する用意もある。実際のミラーリングに際しては各ジャーナルの編集委員会等と個別の調整を実施することになる。

2.2 メタデータ変換とマッチング

ハーベストしたメタデータは DML-JP のプラットフォームである EPrints3.1 へ登録する XML 形式へ変換する。このとき、論文と Math. Reviews の MR 番号、主題分類である MSC との対応をとるために AMS の用意するツール (<http://www.ams.org/mref>) を利用する。MR 番号とのマッチングを取ることで、Math. Reviews の該当するエントリへのリンクを形成することができる。Math. Reviews から DML-JP へのリンク形成は今後の課題である。

2.3 インターフェース

DML-JP 上にはハーベストしたメタデータと Math. Reviews に関連するメタデータとが存在する。機関リポジトリに後者のメタデータを提供するインターフェースを作成した。可能であれば

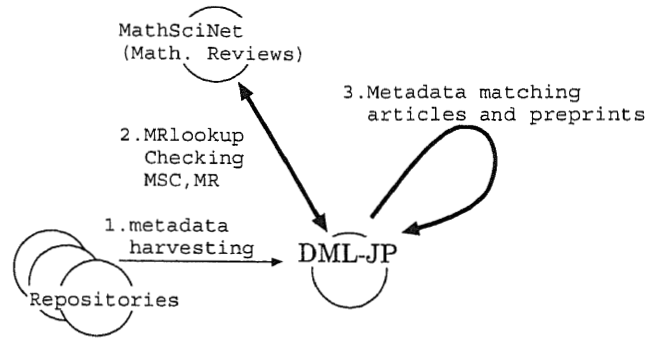


図 2: メタデータ変換とマッチングの概念図

ば、個別に機関リポジトリへ提供したい。当初は web API として設計したが、現在は EPrints 上の SWORD 対応機能を利用して、SWORD 形式で提供する予定である。同時に、主題分類を利用して OAI-ORE の Atom serialization 仕様に準拠した atom 形式 XML を提供している。

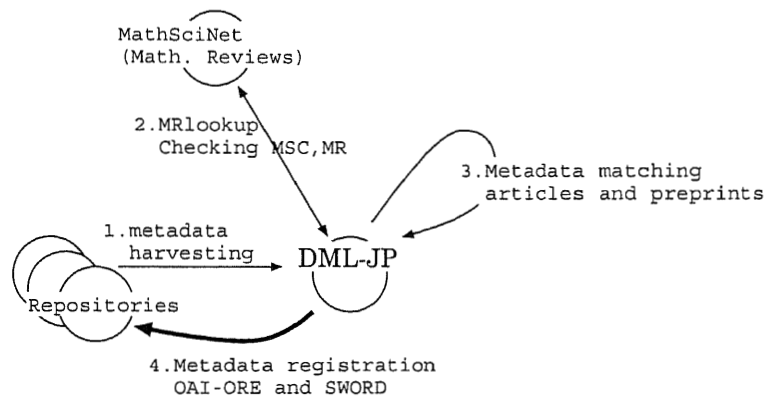


図 3: インターフェース部分の概念図

3 到達点と評価、発展

DML-JP として、日本の数学系ジャーナルに関し MR 収録論文の半数に迫るメタデータを収集した。MR が各論文へ付与する ID と各論文のメタデータとを同定し、これを利用した各レビューへのリンクを形成した。同定に当たっては MR 提供のツールおよび NII 相澤研究室提供のツールを利用した。MR からポータルへのリンクに関しては URL と ID との対を MR へ提供することで実現できるが、各ジャーナルとの調整が残っている。

Digital Mathematics Library 2009 などの国際会議では、分散配置されたデジタルリポジトリから主題ベースのポータルを形成する特徴的な DML として評価された。欧州では国別の DML を統合し EU 全体としての DML を形成するプロジェクトが採択されている。EU の小規模な DML は

日本の数学系ジャーナル数誌に相当する程度でもあり、類似性に注目されている。MRとの関係として、メタデータ利用に関する提案を提出することになっている。MRの保持するメタデータを利用することで、良質なメタデータを確保し、リポジトリに限らず Journal@rchive、J-STAGE 収録のジャーナル等を積極的に利用することが可能になる。

実験的に、国立情報学研究所の IRDB からハーベストした全メタデータについて Math. Reviews とのマッチングを行っている。試験的な結果ではあるが、機関リポジトリに搭載された極めて小規模(論文数 10 程度)のタイトルを把握できている。この結果を DML-JP と同様に EPrints へ搭載した。URL は <http://dmljp2.math.sci.hokudai.ac.jp/view/publication/> である。メタデータからは著者版との区別をつけ難いために、現状の IRDB はタイトルの把握目的として利用するべきと思われる。

arXiv.org などのプレプリントサーバに登録されたプレプリントとの MR および本論文との同定を目指す。プレプリントと本論文との同一性を確保することは、多様なコンテンツの混在する現在において重要な課題である。

フルテキストのアーカイブを実現し、数式 OCR による XHTML 化 [5] などを試行、ディスカッションプラットフォームとしての発展も実装したい。

4 謝辞

DML-JP は 2008 年度の国立情報学研究所 SPARC Japan 事業の支援を受けて構築した。ハーベストしたメタデータと Math. Reviews とのマッチングには国立情報学研究所相澤研究室に協力を受けた。EPrints からの junii2 対応 SWORD 出力は北海道大学附属図書館との協力で実装したものである。

参考文献

- [1] Takao Namiki, Hiraku Kuroda and Shunsuke Naruse, *Experimental DML over digital repositories in Japan*, In proceedings of Towards a Digital Mathematics Library (Ed. Petr Sojka, 2009, Grand Bend, Canada)
- [2] Takao Namiki, *Current status of mathematical publications in Japan*, In proceedings of Towards Digital Mathematics Library pp. 97-102 (Ed. Petr Sojka, July 27th, 2008, Birmingham, UK)
- [3] Takao Namiki, *Status of mathematical publication in Japan: Institutional repositories play an essential role*, In Proceedings of Open Repositories 2008 (March 30th - April 4th, 2008, Southampton, UK)
- [4] Project Euclid <http://projecteuclid.org>
- [5] Infty Project <http://www.sciaccessnet.org/>
- [6] WDML <http://www.mathunion.org/Wdml/>

A DML-JP の対象となるジャーナルリスト

表 1 に主要な数学系ジャーナルのリストを与える。これ以外にも数百のタイトルが存在し、その全貌はわからない。タイトルには Math. Reviews 内の省略形を採用した。発行が続いているタイトルはカレント欄に y で示している。電子版を持つ場合、プラットフォームを示した。ほとんどのタイトルはコミュニティベースのプラットフォームである。SPARC Japan の電子化支援により、大規模なジャーナルは Project Euclid をプラットフォームとして選択している。比較的小規模なジャーナルには機関リポジトリもしくは独自の web ページが多い。CSI 事業による機関リポジトリの浸透と、初期コンテンツとしての紀要登録とから、この傾向は続くと考えられる。

論文数	タイトル	プラットフォーム	カレント
7830	Surikaisekikenkyusho Kokyuroku	IR	y
2753	Proc Japan Acad	Project Euclid	y
2395	Proc Japan Acad Ser A Math Sci	Project Euclid	y
2391	J Math Soc Japan	Project Euclid	y
2357	Tohoku Math J	Project Euclid	y
2079	Nagoya Math J	Project Euclid	y
2010	Osaka J Math	Project Euclid	y
1679	J Math Kyoto Univ	Project Euclid	y
1486	Publ Res Inst Math Sci	Project Euclid	y
1363	Sugaku		y
1242	Hiroshima Math J	Project Euclid	y
1082	Hokkaido Math J	Project Euclid/IR	y
1028	Funkcial Ekvac	J-STAGE	y
972	Kodai Math J	Project Euclid	y
952	Tokyo J Math	Project Euclid	y
915	Tsukuba J Math	IR	y
898	Math J Okayama Univ	IR	y
897	Kodai Math Sem Rep	Project Euclid	n
656	J Fac Sci Univ Tokyo Sect IA Math	IR	n
652	Yokohama Math J	IR	y
626	Mem Fac Sci Kyushu Univ Ser A	J-STAGE	n
547	Math Sem Notes Kobe Univ	Departmental web	n
498	Japan J Math	Springer	y
380	Japan J Indust Appl Math	Project Euclid	y
376	Kobe J Math	Departmental web	y
345	Proc Imp Acad Tokyo	Project Euclid	y
315	Kyushu J Math	J-STAGE	y

表 1: 300 以上の論文を持つ数学系ジャーナルのタイトル