Mathematical journals in Japan and digital mathematics library

Takao Namiki and Hiraku Kuroda
Department of Mathematics,
Hokkaido University,
060-0810 Sapporo, Japan

9 Sep. 2009 RIMS workshop, Kyoto

9400 Nuclear-Phys. B
9131 Pacific-J. Math
8517 Appl. Math. Comput
7865 Math. Z
7491 Theoret. Comput. Sci

About 2,450,000 articles in 12,400 journal titles are indexed in Math. Reviews database from 1940.

- 2,000 titles have more than 100 articles,
- 400 titles have more than 1,000 articles and
- 5 titles have more than 10,000 articles.

Currently 2,700 serials are indexed cover-to-cover.

These titles are essential in mathematical communication and no “core journals” exist in mathematics by that reason. Many titles are based on efforts of community.

Table: Math. Reviews

<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>#</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>9131</td>
<td>Pacific-J. Math</td>
<td>8635</td>
<td>Phys. Lett. A</td>
</tr>
<tr>
<td>7865</td>
<td>Math. Z</td>
<td>7667</td>
<td>RIMS Kokyuroku</td>
</tr>
</tbody>
</table>

Mathematical literatures

“Whole literatures relative to mathematics should be digitized with structure, that is, theorems should be tagged for reference and mathematical expressions should be formed for passing to software”.

(From international workshop “Towards Digital Mathematics Library 2008”, Birmingham, UK)

- 数学の全論文を電子的にアクセス可能にする。
  All mathematical articles should have electronic version.
- 各国でのコミュニティの役割が期待された。
  Community based digitization was expected for each country.
- しかし、日本では多様性が欠けたと思われる。
  Diversity includes difficulty in Japan.
Backgrounds of DML-JP

1. About 70,000 articles in 400 mathematical journals have been published in Japan.¹ (Math. Reviews)
2. Organized digitization activity have been behind compared with other countries.
3. In recent years several digitization activities are established around digital repositories.
4. Environment for metadata based DML was prepared.
5. DML-JP is a metadata based DML.


Digitization activities

**Kobe Group** Funkcialaj Ekvacioj (Division of Function Equation, Math. Soc. Japan) with InftyProject.

**Japan Science and Technology Agency** J-STAGE (platform for online journals for learned society) and Journal@rchive (digitization of selected journals)

**National Institute of Informatics** SPARC Japan (promoting scholarly publishing) from 2003.

**Institutional Repository** About 80 University Library had launched their institutional repositories supported by NII from 2006.

DML-JP, Digital Mathematics Library, Japanese part

DML-JP is supported by SPARC JAPAN and IR project of NII as part of national portal and Mathematical Society Japan.

▶ メタデータベースの DML。Metadata harvesting based DML.
▶ 数学系ジャーナルと紀要についてはタイトルごとにハーベスト。Title based harvesting for math. journals.
▶ αver. http://dmljp.math.sci.hokudai.ac.jp
27 journals are joined in DML-JP.

- Bull. Tokyo Gakugei University Sec. I (IR)
- Bulletin of College of Science the University Ryukyu (IR)
- Hiroshima Math. J. (Euclid)
- Hokkaido Mathematical Journal (IR/Euclid; Infty)
- J. Math. Soc. Japan (Euclid; Infty)
- Journal of Mathematical Sciences, The University of Tokyo (IR)
- Journal of the Faculty of Education, Kagoshima University (IR)
- Journal of the Faculty of Science, Kagoshima University (IR)
- Journal of the Faculty of Science Shinshu University (IR)
- Journal of the Faculty of Science, the University of Tokyo Sect 1 A (IR)
- Journal of the Faculty of Science, Yamagata University (IR)
- Kodai Math. J. (Euclid)
- Nagoya Math. J. (Euclid)
- Nihonkai Mathematical Journal (IR)
- Osaka J. Math. (Euclid)
- Reports of the Faculty of Science and Engineering, Saga University. Mathematics (IR)
- RIMS Kokyuroku (IR; Infty)
- Ryukyu Mathematical Journal (IR)
- The science reports of the Kanazawa University (IR)
- Tohoku Math. J. (Euclid)
- Tokyo J. of Math. (Euclid)
- Tsukuba Journal of Mathematics (IR/Euclid)
Location of mathematics departments where their journals are joined with DML-JP.

**Experimental Part**

- There are huge number of very small journals (10 to 100 articles per title, 200 journal ?).
- We expect that these titles will be digitized in IRs.
- From IRDB in NII, aggregator of institutional repositories in Japan, we harvested full metadata.
- Matching the metadata with MathSci database.

http://dmljp2.math.sci.hokudai.ac.jp/view/publication

**On preprint and author version**

- In digital repository, various preprints and author version should be identified with the original (publisher) version.
- We need the method to identify preprints and the article which has the same contents.
- In the case of multiple platform, the same problem occur.
Once identified, we can describe by OAI-ORE.

---

**Conclusion 1**

- Digital Math. Library 2009 にて、JSTOR の J. Burns は次のよう に語る。「リポジトリとしての JSTOR は citation などの学術情報基盤を形成する。これを拡張し、高度化するのはコミュニティの仕事である。」
- In DML2009, John Burns said: **JSTOR can create the supporting infrastructure for that network e.g. from citations and from some of our other similarity work.**
- But the extension, cleaning and refinement of the network can only be done by a community of experts.
- 研究者コミュニティ、リポジトリ形成、そして、間に立つ情報基盤のプロが必要になってくると考えられる。

---

**Conclusion 2**

- In DML2009, David Ruddy (Project Euclid) said: **Where can we go?**
  - Abandon the notion of a centrally planned, managed, administered, or controlled “library.”
  - Embrace a network paradigm.
  - Define reasonable local and community areas of responsibility and work fostering a network.
- 存在するものをうまく使っていく、リポジトリをうまく使わせるように組織、制度を作ることも重要。
Almost all digital repositories support OAI-PMH as a function of metadata services.

- Open Archives Initiative Protocol for Metadata Harvesting (www.openarchives.org/OAI)
- Provides all metadata contained in a digital repository by XML with REST.

The following is a part of OAI-PMH requests.

- **Identify** provides information of the repository.
- **ListMetadataFormats** provides list of metadata types.
- **GetRecord** provides full record of specified item.
- **ListRecords** provides all metadata.
- **ListSets** provides list of “Set” which classify the contents.

By requesting **ListRecords** repeatedly for a digital repository, we can get all the metadata. Our target repositories are 16 IRs, Project Euclid and arXiv.org.

An example of OAI-PMH GetRecord response


```xml
<OAI-PMH>
  <responseDate>2009-07-02T07:34:42Z</responseDate>
  <GetRecord>
    <record>
      <header>
        <datestamp>2007-05-23</datestamp>
      </header>
      <metadata>
        <oai_dc:dc>
          <dc:title>Absorption problems for quantum walks in one-dimensional spaces determined by a 2 times 2 unitary matrix U on a state space is finite or infinite by using a new path integral approach orthonormal basis P, Q, R and S of the vector space of complex matrices. Our method studied here is a natural extension of classical random walks.</dc:title>
          <dc:description>Comment: 15 pages, small corrections, journal reference added</dc:description>
          <dc:date>2003-02-07</dc:date>
          <dc:type>text</dc:type>
        </oai_dc:dc>
      </metadata>
    </record>
  </GetRecord>
</OAI-PMH>
```
Implementation

After harvesting metadata via OAI-PMH, we made DML-JP by loading them into certain platform.

**Platform**
All metadata is loaded on EPrints 3.1.1.

**Metadata**
harvesting from ProjectEuclid (Cornell) and institutional repositories in Japan

**Metadata** transformation
- from oai_dc to EPrintsXML
- from junii2 to EPrintsXML

**Metadata** merging bibliographic metadata from repositories with mathematical metadata.

---

### Metadata harvesting

- The number of metadata for target articles is more than 30,000.
- About the half of all articles published in Japanese mathematical journals are grasped.
- Two metadata format:
  - oai_dc which Project Euclid provides
  - junii2 format which is standard format for institutional repositories for metadata exchange in Japan.

---

### Example of oai_dc (1/2)

```xml
<record>
<header>
<identifier>oai:CULeuclid:euclid.jmsj/1240435759</identifier>
<datestamp>2009-04-23</datestamp>
<setSpec>jmsj</setSpec>
</header>
<metadata>
<oai_dc>
<dc:title>Minimal 2-regular digraphs with given girth</dc:title>
<dc:creator>BEHZAD, Mehdi</dc:creator>
<dc:subject>05C20</dc:subject>
<dc:subject>05C20</dc:subject>
<dc:publisher>Mathematical Society of Japan</dc:publisher>
```
junii2 and institutional repositories in Japan

**Advantage** Bibliographic element is defined as an entity, which makes it easy to retrieve bibliographic information.

**Difficulty** Some institutional repository does not include journal title in English and even if included the expression does not coincide the expression of Math. Reviews. By that reason it is relatively hard to retrieve MR number and MSC from Math. Reviews database.

After metadata harvesting, the two metadata formats were transformed into EPrints XML format.

It is easy because EPrintsXML define all entities required oai_dc and junii2.

For mathematical metadata, MSC and MR, we added mr, msc_p and msc fields to the set of EPrintsXML.

Example of EPrintsXML (1/2)

```xml
<?xml version="1.0" encoding="utf-8" ?>
<eprints>
  <eprint xmlns="http://eprints.org/ep2/data/2.0">
    <rev_number>1</rev_number>
    <eprint_status>archive</eprint_status>
    <userid>1</userid>
    <metadata_visibility>show</metadata_visibility>
    <type>article</type>
    <ispublished>pub</ispublished>
    <subjects>
      <item>20-xx</item>
      <item>QA</item>
    </subjects>
    <refereed>TRUE</refereed>
    <full_text_status>public</full_text_status>
    <date_type>published</date_type>
    <datestamp>2007-08-01T01:50:05Z</datestamp>
  </eprint>
</eprints>
```

Example of EPrintsXML (2/2)

```xml
<title>Note on the Schur multiplier of a certain semidirect product</title>
<creators_name>
  <item>
    <family>Horie</family>
    <given>Mitsuko</given>
  </item>
</creators_name>
<official_url>http://hdl.handle.net/10083/839</official_url>
<pagerange>85-88</pagerange>
<publisher>Ochanomizu Univeristy</publisher>
<msc_p>20J06</msc_p>
<msc>20C25</msc>
<mr>1317509</mr>
<related_url>
  <item>
  </item>
</related_url>
</eprints>
```
Example: dmljp.math.sci.hokudai.ac.jp/32786/.

IR Author: Maeda, Masao
IR Title: The four-or-more Vertex Theorems in 2-dimensional Space Forms
IR Official URL: http://hdl.handle.net/10131/1069
MR MSC Primary: 53A35, 53A, 53
MR MSC Secondary: 53A04, 53A, 53
MR Math. Reviews ID: 1710269

Though this journal is so small and interdisciplinary that only this article is reviewed and indexed in Math. Reviews, you can find in the review URL that this article was cited from a review article in the field.

MR gateway

Because several journals have difficulty to get MR number for their articles, we prepare interface between the articles and MR numbers as follows.

▶ sparcl.math.sci.hokudai.ac.jp/mrgw.cgi?mr=818212

By the usage the result is bibliographic information of the specified article and link to the original repository.

By MR gateway we can resolve MR number to original URL, however, an essential solution is to load MR number and MSC information into original repositories. In the context we have two methods:

▶ SWORD protocol
▶ OAI-ORE
We intend to establish resource finding and exchange schema between digital repositories by the implementation, which is merely experimental phase.

The following is a part of an example of ORE Atom serialization.

```xml
<?xml version="1.0" encoding="utf-8" ?>
  <mets:metsHdr CREATEDATE="2009-06-30T05:48:08Z">
    <mets:agent TYPE="ORGANIZATION" ROLE="CUSTODIAN">
      <mets:name>Hokkaido Mathematical Journal</mets:name>
    </mets:agent>
  </mets:metsHdr>
  <mets:dmdSec ID="DMD_oai_hmj.math.sci.hokudai.ac.jp_6_junii2:junii2">
    <mets:mdWrap OTHERMDTYPE="JUNII2" MDTYPE="OTHER">
      <juni2:metadata>
        <juni2:title>Real moduli in local classification</juni2:title>
        <juni2:creator>MORMUL, Piotr</juni2:creator>
        <juni2:subject>Goursat flag, singularity, local c</juni2:subject>
        <juni2:subject>58A30(MSC2000)</juni2:subject>
        <juni2:subject>58A17(MSC2000)</juni2:subject>
        <juni2:subject>QA Mathematics</juni2:subject>
        <juni2:description>pages: 1</juni2:description>
        <juni2:format>application/pdf</juni2:format>
        <juni2:URI>http://hmj2.math.sci.hokudai.ac.jp/6/</juni2:URI>
      </juni2:metadata>
    </mets:mdWrap>
  </mets:dmdSec>
</mets:mets>
```

SWORD is an inter repository interface based on Atom Publishing Protocol.

- EPrints3 provides exporting facility for METS suitable for SWORD protocol.
- By the facility we are planning to give the mathematical metadata into original institutional repositories.
Acknowledgement

This work is mainly supported by SPARC JAPAN [3] from Apr. 2008 to Mar. 2009, and had been supported by Department of mathematics, Hokkaido University under governmental funding of 21st century Center of Excellence. Mathematical Society of Japan also supports this activity.


SPARC JAPAN, National Institute of Informatics http://www.nii.ac.jp/sparc/

Cyber Science Infrastructure, National Institute of Informatics http://www.nii.ac.jp/irp/

ORE Specification http://www.openarchives.org/ore/

EPrints http://www.eprints.org/

InftyReader http://www.sciaccessnet.org/

Ayman Ferahat, Thomas Lofaro, Joel C. Miller, Gregory Rae, and Lesley A. Ward., *Authority rankings from HITS, PageRank, and SALSA: Existence, uniqueness, and effect of initialization.*, SIAM Journal on Scientific Computing. 27 (4) 1181-201