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

Mathematical literatures

数学の論文とよべるものは、Math. Reviews database に依拠すると次のような規模と考えられる。
Scale of journal articles in mathematics.

- 1940 年以降、約 2,450,000 論文が 12,400 ほどのタイトルから発表されている。
  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.

非常に多様であって、コアジャーナルと呼べる存在がない。コミュニティベースの出版形態も多い。
These titles are essential in mathematical communication and no "core jouornals" exist in mathematics by that reason. Many titles are based on efforts of community.

Table: Math. Reviews に収録されたジャーナルから、論文数で上位 20 件を抜き出した。青は数学のジャーナルと考えられるタイトル。近隣分野との相互作用も重要な要素。 Top 20 in number of articles from Math. Reviews indexed journals. Blue title: Mathematics, Black title: other fields: half of these titles are non-mathematical.

#	Title	#	Title
24296	Proc. Amer. Math. Soc	17720	J. Phys. A
17199	J. Math. Anal. Appl	14366	Trans. Amer. Math. Soc
12976	J. Math. Phys	9849	Phys. Rev. D (3)
9805	J. Algebra	9792	Phys. Lett. B
9709	Dokl. Akad. Nauk-SSSR	9649	Discrete-Math
9400	Nuclear-Phys. B	9289	C. R. Acad. Sci. Paris-Ser.
9131	Pacific-J. Math	8635	Phys. Lett. A
8517	Appl. Math. Comput	7929	Linear-Algebra-Appl
7865	Math. Z	7667	RIMS Kokyuroku
7491	Theoret. Comput. Sci	7450	IEEE-Trans. Automat. Cor

Concept of Digital Mathematics Library

"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、日本では多様性が仇になったと思われる。
  Diversity includes difficulty in Japan.
- デジタルリポジトリの役割。
  Role of digital repositories.

Digital Mathematics Library (by T.Bouche, NUMDAM)

US	JSTOR (260,000 items), project Euclid (100,000)
Asia	DML-JP (30,000 items), China ??
Europe	EuDML? (190,000 items)
Germany	ERAM/JFM, GDZ, ELibM (85,000 items)
France	Gallica-Math, NUMDAM, CEDRAM, TEL (50,000)
Poland	ICM/BWM (13,000 items)
Portugal	SPM/BNP (2,000 items)
Spain	DML-E (5,000 items)
Czech	DML-CZ (11,000 items)
Russia	RusDML (13,000 items)
Bulgaria	BulDML (2,500 items)

Commercial base: 700,000 items? Small/medium CUP 20 journals, OUP 30, Hindawi 18, WdG 13, Wiley 42, T&F 58. . . Elsevier 4 journals in NUMDAM, 63 in Backfiles, 100 alive (320,000 items) Springer 14 journals in GDZ, 1+2 in NUMDAM, 120 in Online Archives, 179 alive (300,000 items)

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.

[1] Takao Namiki, *Current status of mathematical publications in Japan*, In proceedings of Towards Digital Mathematics Library 2008 pp. 97–102 (Ed. Petr Sojka, July 27th, 2008, Birmingham, UK)

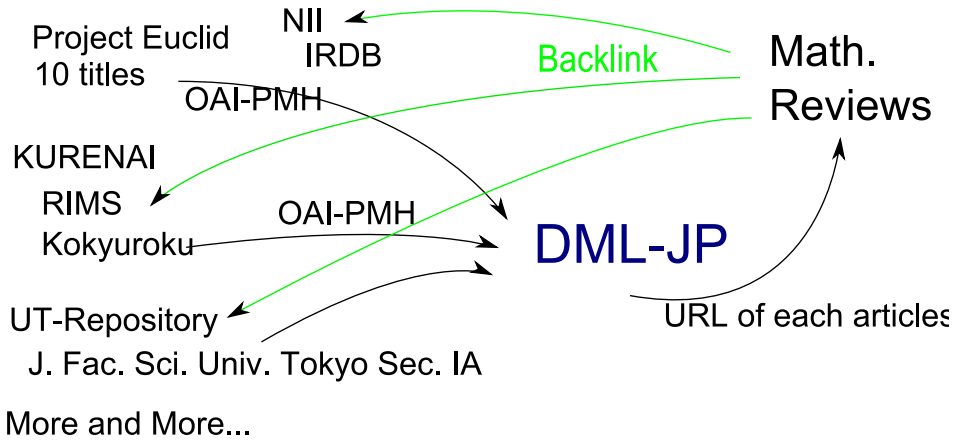
Digitization activities

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.
- α*ver*. `http://dmljp.math.sci.hokudai.ac.jp`
- β*ver*. `http://sparc1.math.sci.hokudai.ac.jp/dmljp/`



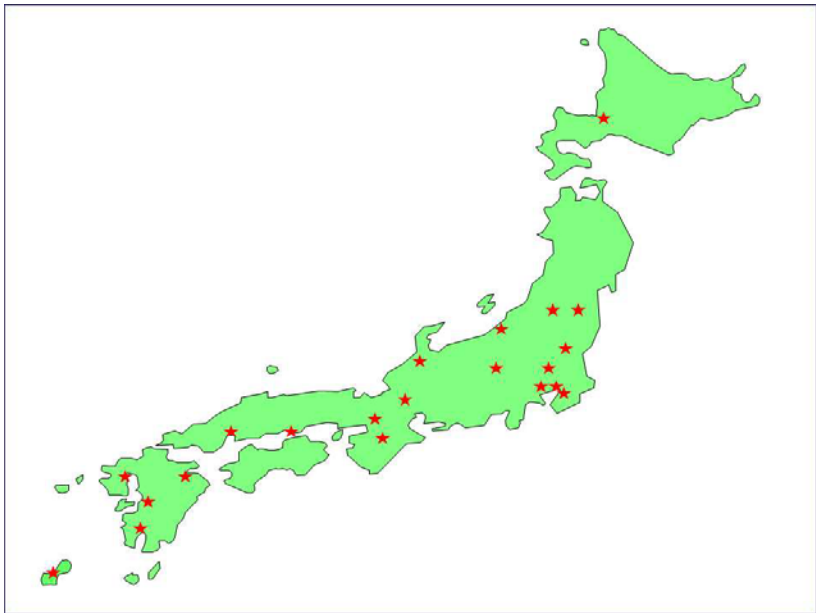
27 journals are joined in DML-JP.

- 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)
- Japan J. Indust. Appl. Math. (Euclid)
- 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)
- Nat. Sci. J. Fac. Educ. Hum. Sci. Yokohama National University Sec. I (IR)
- Natur. Sci. Report. Ochanomizu. Univ. (IR)
- Nihonkai Mathematical Journal (IR)
- Osaka J. Math. (Euclid)
- Proc. Japan Acad. Ser. A Math. Sci. (Euclid)
- Publ. Res. Inst. Math. Sci. (Euclid)

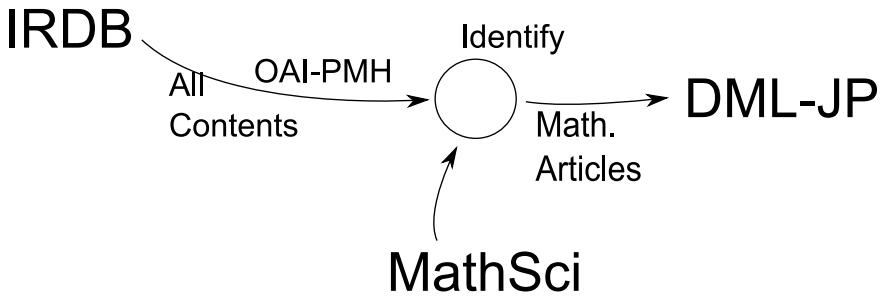
- Reports of the Faculty of Science and Engineering, Saga University. Mathematics (IR)
- RIMS Kokyuroku (IR; Infty)
- Ryukyu Mathematical Journal (IR)
- Sci. Rep. Yokohama National University Sec. I (IR)
- The science reports of the Kanazawa University (IR)
- Tohoku Math. J. (Euclid)
- Tokyo J. of Math. (Euclid)
- Tsukuba Journal of Mathematics (IR/Euclid)

Location of mathematics departments where their journals are joined with DML-JP.



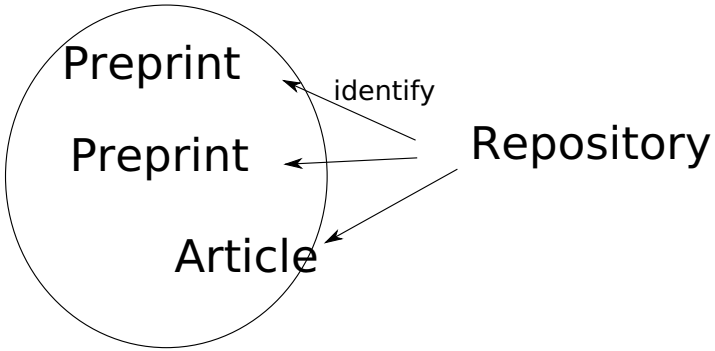
Experimental Part

- 多数の極小規模紀要の存在。
  There are huge number of very small journals (10 to 100 articles per title, 200 journal ?).
- この種のジャーナルは institutional repository へ入るのではないかと 期待。
  We expect that these titles will be digitized in IRs.
- 包括的なハーベスティングに NII の IRDB を利用する。
  From IRDB in NII, aggregator of institutional repositories in Japan, we harvested full metadata.
- IRDB と MathSci との同定処理から" 数学" の論文を抽出。
  Matching the metadata with MathSci database.
- `http://dmljp2.math.sci.hokudai.ac.jp/view/publication`



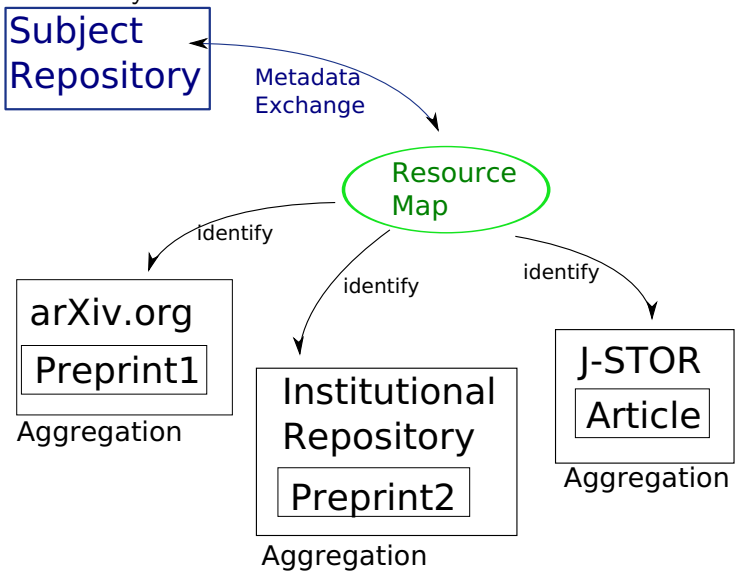
プレプリント、著者版の取扱い On preprint and author version

- デジタルリポジトリを巡っては、プレプリント、著者版などの同定問題が出現する。
  In digital repository, various preprints and author version should be identified with the original (publisher) version.
- プレプリントと対応する論文の組を同定する手段が欲しい。
  We need the method to identify preprints and the article which has the same contents.
- 複数のプラットフォームを利用している場合も同様。(cf. JMSJ, J-STAGE+Euclid)
  In the case of multiple platform, the same problem occur.



OAI-ORE realization

同定できれば OAI-ORE で記述できる。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 communty 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。

OAI-PMH

Almost all digital repositories supports 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.
- Simple to use in HTTP GET request, `http://baseurl/cgi?verb=`*verb*.

OAI-PMH request

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	provedes 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

```
http://export.arxiv.org/oai2?verb=GetRecord&
metadataPrefix=oai_dc&
identifier=oai:arXiv.org:quant-ph/0208122
```

```
<OAI-PMH>
<responseDate>2009-07-02T07:34:42Z</responseDate>
<request verb="GetRecord" identifier="oai:arXiv.org:quant-p
<GetRecord>
<record>
  <header>
    <identifier>oai:arXiv.org:quant-ph/0208122</identifier>
    <datestamp>2007-05-23</datestamp>
    <setSpec>physics:quant-ph</setSpec>
  </header>
  <metadata>
    <oai_dc:dc>
      <dc:title>Absorption problems for quantum walks in on
      <dc:creator>Namiki, Takao</dc:creator>
      <dc:subject>Quantum Physics</dc:subject>
```

```
 <dc:description>
  This paper treats absorption problems for the one-dimensi
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 comp
matrices. Our method studied here is a natural extension of
classical random walk.
</dc:description>
<dc:description>
Comment: 15 pages, small corrections, journal reference add
</dc:description>
<dc:date>2003-02-07</dc:date>
<dc:type>text</dc:type>
<dc:identifier>http://arxiv.org/abs/quant-ph/0208122</dc:id
<dc:identifier>
J. Phys. A: Math. Gen., Vol. 36, No.1, pp.241-253 (2003)
</dc:identifier>
</oai_dc:dc>
</metadata>
```

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

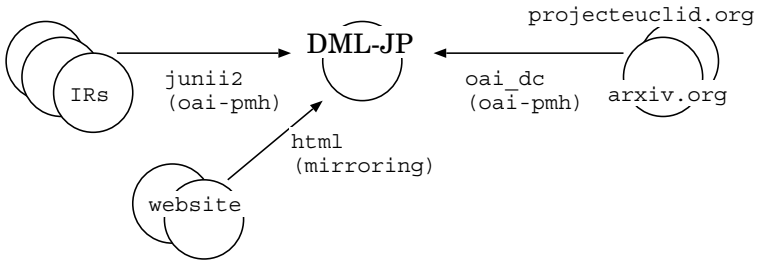


Figure: Metadata harvesting

- Harvesting 16 institutional repositories, Project Euclid and arXiv.org.
- 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)

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

Example of oai_dc (2/2)

```
  <dc:date>1973-01</dc:date>
  <dc:type>Text</dc:type>
  <dc:format>application/pdf</dc:format>
  <dc:identifier>
   http://projecteuclid.org/euclid.jmsj/1240435759
  </dc:identifier>
  <dc:identifier>
   J. Math. Soc. Japan 25, no. 1 (1973), 1-6
  </dc:identifier>
  <dc:identifier>doi:10.2969/jmsj/02510001</dc:identifier>
  <dc:language>en</dc:language>
  <dc:rights>
   Copyright 1973 Mathematical Society of Japan
  </dc:rights>
</oai_dc:dc>
</metadata>
</record>
```

oai_dc

Advantage Simple and standard for OAI-PMH compliant digital repositories.

Difficulty Description of bibliographic information in `dc:identifier`.

Example of junii2 (1/2)

```
<record>
<header>
<identifier>oai:teapot.lib.ocha.ac.jp:10083/843</identifier
<datestamp>2007-07-02T06:30:00Z</datestamp>
<setSpec>hdl_10083_792</setSpec>
</header>
<metadata>
<meta xmlns="http://ju.nii.ac.jp/junii2">
<title>
 CONDITIONALLY TRIMMED SUMS FOR INDEPENDENT RANDOM VARIABLE
</title>
<creator>KASAHARA, Yuji</creator>
<NDC>400</NDC>
<publisher>Ochanomizu University</publisher>
<type>Article</type>
<NIItype>Departmental Bulletin Paper</NIItype>
<format>application/pdf</format>
<format>191755 bytes</format>
```

Example of junii2 (2/2)

```
<URI>http://hdl.handle.net/10083/843</URI>
<fullTextURL>
http://teapot.lib.ocha.ac.jp/ocha/bitstream/10083/843/1/KJ(
</fullTextURL>
<issn>00298190</issn>
<NCID>AN00033958</NCID>
<jtitle>Natur. Sci. Rep. Ochanomizu Univ.</jtitle>
<volume>46</volume>
<issue>2</issue>
<spage>9</spage>
<epage>12</epage>
<dateofissued>1995-12-30</dateofissued>
</meta>
</metadata></record>
```

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 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>
  <publication>Natur. Sci. Report. Ochanomizu. Univ.</publi
  <datestamp>2007-08-01T01:50:05Z</datestamp>
```

Example of EPrintsXML (2/2)

```
  <title>
  Note on the Schur multiplier of a certain semidirect proc
  </title>;
  <creators_name><item><family>Horie</family>
    <given>Mitsuko</given></item></creators_name>
  <official_url>http://hdl.handle.net/10083/839</official_u
  <pagerange>85-88</pagerange>
  <volume>45</volume>
  <date>1994-12-15</date>
  <publisher>Ochanomizu Univeristy</publisher>
  <msc_p>20J06</msc_p>
  <msc><item>20C25</item></msc>
  <mr>1317509</mr>
  <related_url><item>
   <url>http://www.ams.org/mathscinet-getitem?mr=1317509</u
   <type>MathSciNet</type></item></related_url>
  </eprint>
</eprints>
```

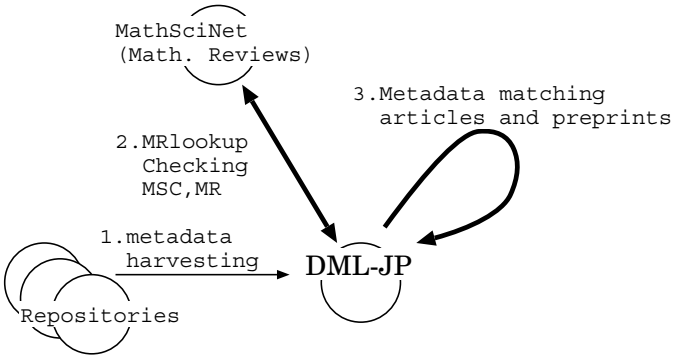


Figure: Metadata processing

Example: dmljp.math.sci.hokudai.ac.jp/32786/.

```
IR Author: Maeda, Masao
IR Title: The four-or-more Vertex Theorems in 2-dimensior
   Space Forms
IR Citation: Nat. Sci. J. Fac. Educ. Hum. Sci. Yokohama N
             University Sec. I, 1 (1998) . pp. 43-46.
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
MR Review URL: http://www.ams.org/mathscinet-getitem?mr=1
```

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.

- `sparc1.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.

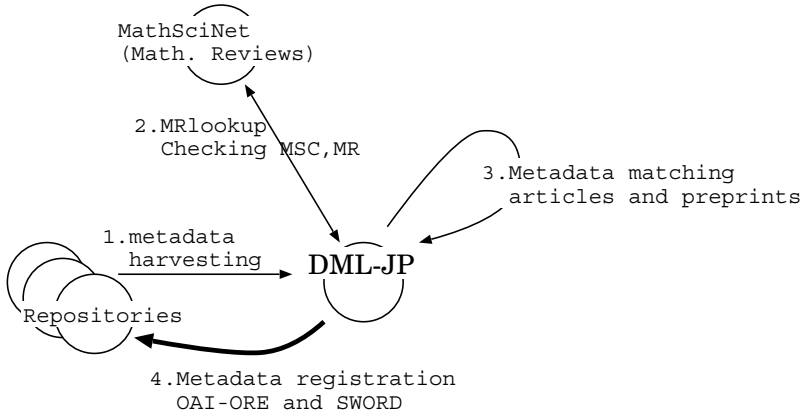


Figure: Metadata registration

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

```
<!-- Aggregated Resources -->
<atom:link href='http://projecteuclid.org/euclid.kmj/11388
  title='A remark on derived spaces'
  rel='http://www.openarchives.org/ore/terms/aggregates' /
<atom:link href='http://projecteuclid.org/euclid.tmj/11921
  title='Spectral synthesis in the Fourier algebra and the
     Varopoulos algebra'
  rel='http://www.openarchives.org/ore/terms/aggregates' /
```

SWORD

- 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.

```
<?xml version="1.0" encoding="utf-8" ?>
<mets:mets>
<mets:metsHdr CREATEDATE="2009-06-30T05:48:08Z">
<mets:agent TYPE="ORGANIZATION" ROLE="CUSTODIAN">
<mets:name>Hokkaido Mathematical Journal</mets:name></mets:
<mets:dmdSec ID="DMD_oai_hmj.math.sci.hokudai.ac.jp_6_junii
<mets:mdWrap OTHERMDTYPE="JUNII2" MDTYPE="OTHER">
<mets:xmlData><junii2:junii2>
        <junii2:title>Real moduli in local classification o
        <junii2:creator>MORMUL, Piotr</junii2:creator>
        <junii2:subject>Goursat flag, singularity, local cl
        <junii2:subject>58A30(MSC2000)</junii2:subject>
        <junii2:subject>58A17(MSC2000)</junii2:subject>
        <junii2:LCC>QA Mathematics</junii2:LCC>
        <junii2:description>pages: 1</junii2:description>
        <junii2:type>article</junii2:type>
        <junii2:NIItype>Journal Article</junii2:NIItype>
        <junii2:format>application/pdf</junii2:format>
        <junii2:URI>http://hmj2.math.sci.hokudai.ac.jp/6/</
```

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.

- Takao Namiki, *Status of mathematical publication in Japan: Institutional repositories play an essential role*, In Open Repositories 2009 abstracts. (2009)
- DML-JP `http://dmljp.math.sci.hokudai.ac.jp/`
- 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