

日本の数学関係ジャーナルの状況とポータルとしての仮想 DML の構築

行木孝夫

nami@math.sci.hokudai.ac.jp,

黒田拓

hiraku@math.sci.hokudai.ac.jp

数学ソフトウェアとフリードキュメント 08

2009年3月25日

東京大学数理科学研究科

Knoppix/Math 2009 上の DML-JP

DML-JPとは、Mathematical Reviews/MathSciNet に収録されている日本の数学系ジャーナル、紀要に掲載された論文をを可能な限り集約したポータルです。MathSciNet 収録誌に対象を限らない場合に JDML と呼ぶことにします。この様なポータルの

- ▶ 実装の意義と
- ▶ 実装の手段を

見ていきます。

現在、下記の URL が稼働しています。

- ▶ `dml.jp.math.sci.hokudai.ac.jp(alpha)`
- ▶ `sparcl.math.sci.hokudai.ac.jp/dml.jp/(beta)`
- ▶ `jdml.math.or.jp/jdml/(static)`

DML-JP は、2008 年度国立情報学研究所国際学術情報流通基盤整備事業 (SPARC Japan) の支援を受けました。

ポータルの意義

ポータル形成に当たっては、次のような問題点を思いつきます。

- ▶ そもそも、**MathSciNet** に日本の数学系ジャーナル、紀要は何誌ほど収録されているのか。
- ▶ 論文数はどれだけか。
- ▶ それらのジャーナルや紀要、掲載された論文は重要なのか。

次のスライドへ移る前に少し考えてみます。

MathSciNet 収録誌の数

- ▶ 誠実な解答は、「わからない」。既に失われた書誌が多数存在する。
- ▶ それでも、300 から 400 タイトルが **Mathematical Reviews** に存在することは確実である。
- ▶ ところで、論文数 “1” から “10” という 紀要などは、どう扱うべきだろうか。

MathSciNet 収録論文数

- ▶ 収録タイトルがわかれば数え上げるだけであるが、
- ▶ 何をもって「数学」の論文とみなすべきだろうか。
- ▶ 物理学会誌などはどうするか。
- ▶ 物理系含めて 70,000 論文程度と評価できる。
- ▶ MathSciNet 収録文献はおよそ 200 万件ほど。3%前後と考えられる。
- ▶ ただし、2000 年以降に限れば 1%を割っている。

ポータル の 意義

- ▶ そもそも、日本のジャーナルや紀要に投稿しているのか？
- ▶ それはともかく、どんな意義があるだろうか。

日本のジャーナル、紀要のパフォーマンス

Table: Math. Reviews/MathSciNet 収録論文に対する分野別シェア

論文数比	論文数/総数	研究分野
10.62	1923/18103	57 多様体と胞複体
10.00	1852/18506	32 複素多変数関数と解析空間
9.48	545/5748	31 ポテンシャル論
9.46	1048/11077	55 代数的位相幾何学
9.20	1902/20655	14 代数幾何学
8.15	3307/40538	53 微分幾何学
7.68	875/11392	13 可換環と可換代数
7.45	525/7041	12 体論と多項式
6.58	2301/34968	11 数論
6.25	734/11742	22 位相群, リ一群
5.84	1922/32891	30 複素一変数関数
5.44	1305/23970	16 結合的環と代数

日本の研究機関からの論文数が多い分野

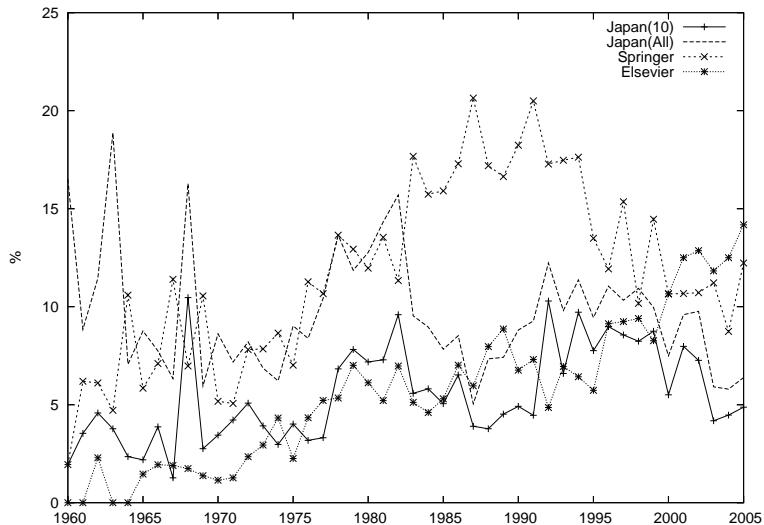
Table: Math. Reviews/MathSciNet 1980 年以降の収録論文に対する分野別シェア

論文数比	論文数/総数	研究分野
16.53	2040.45/12342	57 多様体と胞複体
13.23	2226.33/16826	14 代数幾何学
13.03	766.00/5879	55 代数的位相幾何学
12.31	1614.62/13116	32 複素多変数関数と解析空間
12.19	363.67/2983	31 ポテンシャル論
10.43	2625.82/25182	53 微分幾何学
9.67	783.08/8094	13 可換環と可換代数
9.32	3276.82/35156	11 数論
8.06	4924.28/61097	35 偏微分方程式
7.96	1367.30/17187	30 複素一変数関数
7.92	773.23/9764	17 非結合的環と代数
7.57	527.00/6963	22 位相群, リ一群
7.33	2625.70/35799	58 大域解析, 多様体上の解析

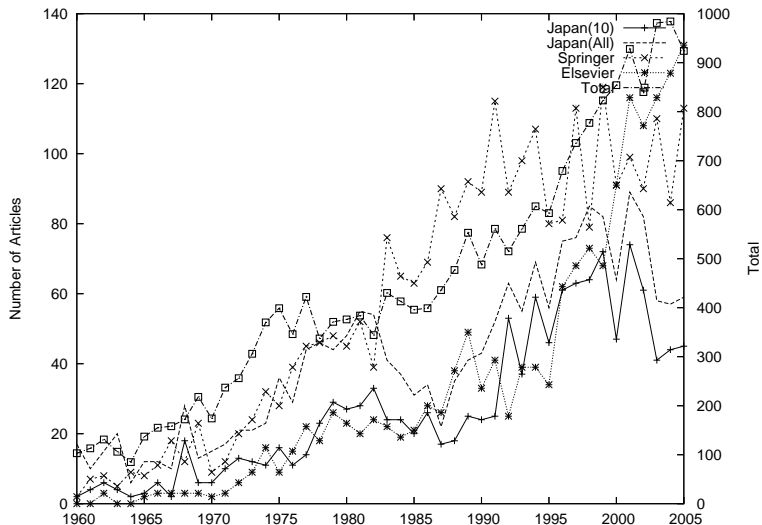
特定分野で時系列を見る

- ▶ 代数幾何学
- ▶ 微分幾何学

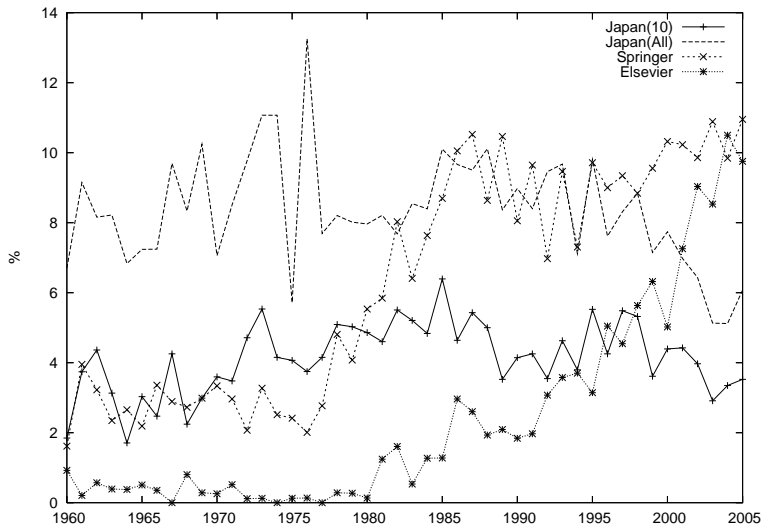
代数幾何学 (1916/19744)



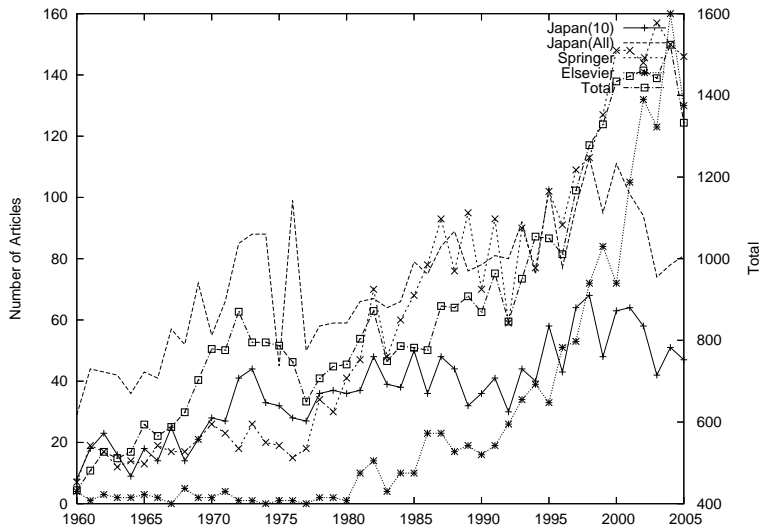
代数幾何学 (総数)



微分幾何学 (3210/39080)



微分幾何学 (総数)



ポータルの意義はあるか

- ▶ 研究分野を限定すれば、一定の位置を占めているものである。
- ▶ 過去の蓄積を維持することは極めて重要なものである。

実装の手段について

- ▶ 対象とする論文は **Digital Repository** 上に分散して存在しているものとする。
- ▶ ポータルとして利用するメタデータを集めなければならない。
 - ▶ メタデータをどのように収集するか。人力ではなく、単純なクロールでもない。
 - ▶ その後、メタデータをどのように利用するか。できれば、**MSC** や **MR** を付与しておきたい。

メタデータ収集手段の背景

- ▶ 学術コミュニケーションに関わる世界的な環境の変化
 - ▶ Serials Crisis and SPARC movement
 - ▶ Open Access
 - ▶ Institutional Repositories (University Library)
 - ▶ Subject based Digital Repositories (arXiv.org, Project Euclid, PubMedCentral, RePEC, ...)
- ▶ 基盤技術の普及
 - ▶ Repository Software (EPrints, DSpace, Fedora, Xoonips,...)
 - ▶ OAI-PMH
- ▶ 結果として、日本では大学紀要のデジタルイズが大学図書館を中心に進展した。

オープンアクセスということであれば一種のフリードキュメントとできるのではないだろうか。

Open Archives Initiative Protocol for Metadata Harvesting

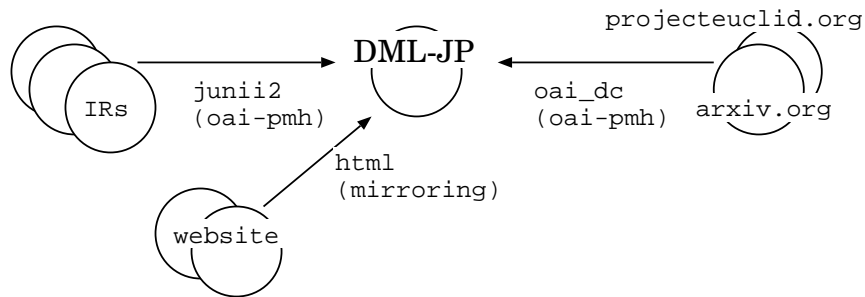
- ▶ プラットフォーム上の全レコードを提供するインターフェース
- ▶ HTTP 上の REST 構造 — BaseURL+parameter
- ▶ `http://baseurl/basecgi?verb=param`

verb	機能
Identify	サーバ情報の出力
ListSets	分類情報の出力
GetRecord	1 エントリーのメタデータ出力
ListRecords	全エントリーのメタデータ出力 (区切りあり)

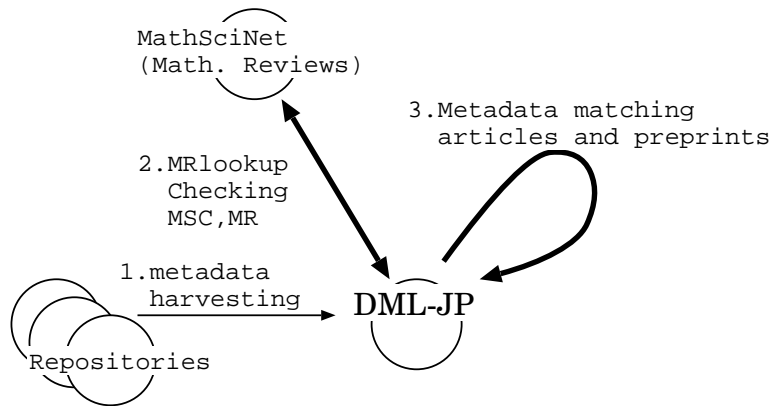
実装

1. OAI-PMH によるメタデータ収集
 - ▶ 14 repositories, 28 journals.
2. XML を変換してサーバソフトウェアへ搭載
 - ▶ EPrints 3.1.1 — repository software
 - ▶ MR code, MSC の付与
3. メタデータのアップデートと搭載

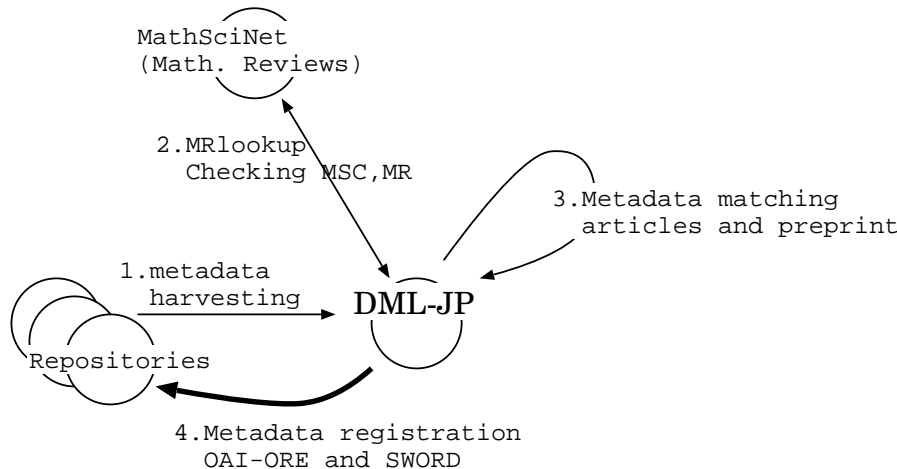
メタデータハーベスティング



MR code, MSC の付与、DML-JP サーバへの登録



Digital Repository へのメタデータインターフェース



収集したメタデータをどのように利用するか。

- ▶ MathSciNet ゲートウェイとして
 - ▶ <http://baseurl/mrgw.cgi?mr=mr-code>
- ▶ フルテキスト 収集、OCR、数式検索
- ▶ マルチメディア対応
 - ▶ ビデオアーカイブ (東大、数学会等)
 - ▶ 貴重資料: 1955 年の日光シンポジウム音声テープ (オープンリール)、谷山・志村・Weil 予想提出三名の講演記録
 - ▶ どう扱い保存していくか
- ▶ インターフェース

今後

- ▶ 例えば、arxiv.org のプレプリントとの対応をとる。
- ▶ [arXiv/math](https://arxiv.org/math) のメタデータを載せる。
- ▶ 機関リポジトリに入っている **Math. Reviews** 収録論文を日本のジャーナルに限らず収録する。
- ▶ **Math. Reviews** との同定の精密化。

- ▶ 2005 年度より RIMS 研究集会「 紀要の電子化と 周辺の話題」を開いてきました。
- ▶ 今年は 9/8-9 で「 数学における デジタルライブラリ 構築へ向けて」として開催します。

付録: Journal titles(1/5)

Total	Title	Platform	Status
7830	<u>Surikaisekikenkyusho Kokyuroku</u>	IR	y
2753	Proc Japan Acad	Euclid	y/1
2395	Proc Japan Acad Ser A Math Sci	Euclid	y
2391	<u>J Math Soc Japan</u>	Euclid	y
2357	<u>Tohoku Math J</u>	Euclid	y
2079	Nagoya Math J	Euclid	y
2010	Osaka J Math	Euclid	y
1679	J Math Kyoto Univ	()	y
1486	Publ Res Inst Math Sci	Euclid	y
1363	Sugaku		y
1242	Hiroshima Math J	Euclid	y
1082	<u>Hokkaido Math J</u>	IR	y
1028	<u>Funkcial Ekvac</u>	JST/web	y

Journal titles(2/5)

Total	Title	Platform	Status
972	Kodai Math J	Euclid	y
952	Tokyo J Math	Euclid	y
915	Tsukuba J Math	IR	y
898	Math J Okayama Univ	IR	y
897	Kodai Math Sem Rep	Euclid	n
656	J Fac Sci Univ Tokyo Sect IA	IR	n
652	<u>Yokohama Math J</u>	(IR)	y
626	Mem Fac Sci Kyushu Univ Ser A	J-STAGE/IR	n
547	Math Sem Notes Kobe Univ	web	n
498	Japan J Math	(Springer)	y
380	Japan J Indust Appl Math	Euclid	y
376	Kobe J Math	web	y
345	Proc Imp Acad Tokyo	Euclid	n

Journal titles(3/5)

Total	Title	Platform	Status
315	Kyushu J Math	J-STAGE/IR	y
297	J Math Sci Univ Tokyo	IR	y
271	<i>Sci Rep Tokyo Kyoiku Daigaku</i>		n
262	<u>Nihonkai Math J</u>	web/IR/(Euclid)	y
244	SUT J Math	web	y
241	Osaka Math J	Euclid	y
219	Natur Sci Rep Ochanomizu Univ	IR	y
218	Mem Fac Sci Kochi Univ Ser A		y
211	Math J Toyama Univ		y
206	J Fac Sci Univ Tokyo Sect I	IR	n
190	Mem Coll Sci Univ Kyoto Ser A	()	n
184	Sugaku Expositions		n
182	<i>Bull Fac Sci Ibaraki Univ Ser A</i>	IR	y
171	J Sci Hiroshima Univ Ser A I	Euclid	n
170	J Sci Hiroshima Univ Ser A	Euclid	n

Journal titles(4/5)

Total	Title	Platform	Status
172	<u>J Fac Sci Hokkaido Univ Ser I</u>	IR	n
147	Math Rep Toyama Univ		n
144	Proc Phys Math Soc Japan	(J-STAGE)	n
143	Bull Yamagata Univ Natur Sci	IR	y
142	Tech Rep Osaka Univ		n
140	J Math Tokushima Univ		-
137	J Fac Sci Shinshu Univ	IR	y
133	Interdiscip Inform Sci	J-STAGE/IR	y
128	<i>Bull Fukuoka Univ Ed</i>		-
125	J Hokkaido Univ Ed Sect		-
120	<u>Sci Rep Niigata Univ Ser A</u>	()	-
118	Bull Nagoya Inst Tech		-
116	Mem Fac Engrg Kyoto Univ		-

Journal titles (5/5)

Total	Title	Platform	Status
116	Bull Tokyo Gakugei Univ	IR	y
110	Sci Rep Kanazawa Univ	IR	y
109	Sci Papers College Gen Ed Univ Tokyo	IR	-
109	Bull Kyushu Inst Tech Math Natur Sci		-
107	Ryukyu Math J	IR	y
105	Proc Fac Sci Tokai Univ		-
101	Saitama Math J		y
99	Sci Rep Kagoshima Univ	IR	y
99	<i>J College Engrg Nihon Univ Ser B</i>		-
96	Math Rep Kyushu Univ		-
93	Mem Fac Sci Shimane Univ	web	-
91	Kumamoto J Sci Ser A		-