Network of Digital Repository in mathematics community

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

7 Sep. 2009 DRF technical workshop, Kyoto

数学の成果は定理と証明の組/"Results" in mathematics is pair of theorem and proof.

証明されていない定理を見つけたら（予想したら）「問題」とか「予想」とよぶ。

When we find a theorem unproved, we call it "problem" or "conjecture" and when we prove it we write a paper. The following example is extremely special case.

Theorem (Fermat's last theorem)

For all $n \in \mathbb{N}$ $(n > 2)$ *there does not exist* $x, y, z \in \mathbb{N}$ *which satisfy the following equation:*

$$x^n + y^n = z^n.$$

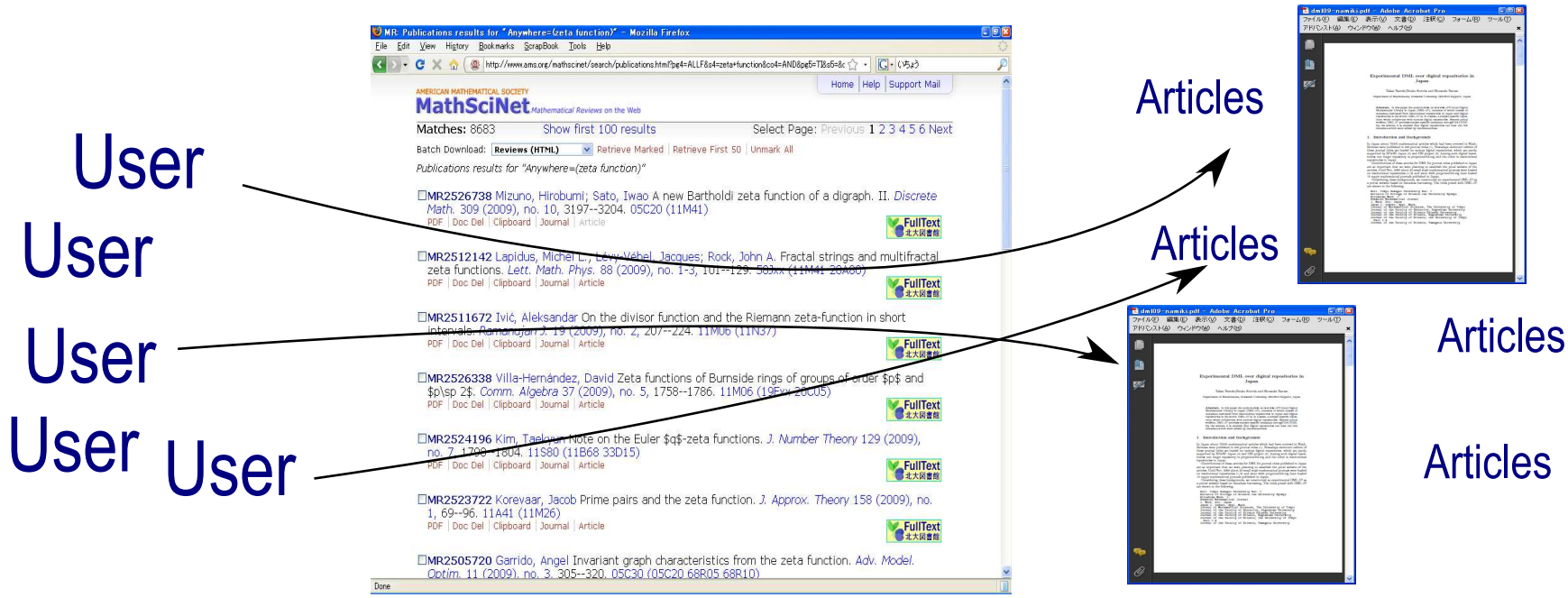
Proof.

We have no time and space to describe the proof... □

Math. Reviews as portal.

主に使う 書誌情報データベースは Math. Reviews であって、ほぼ全ての数学関係論文を網羅していると考えられる。

To find proved theorems, we use Math. Reviews database.



Traditional Mathematical Communication

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.

- About 2,450,000 articles in 12,400 journal titles are indexed in Math. Reviews database from 1940.
- 3,000 titles have more than 10 articles,
- 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	Surikaisekikenkyusho-Koky
7491	Theoret. Comput. Sci	7450	IEEE-Trans. Automat. Cor

Table: 日本発のジャーナルを挙げる。およそ 300 タイトル。プラットフォームは多彩。研究分野によっては 10%程度のシェアがある。
Mathematical journals published in Japan.

Total	Title	Platform	Current
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		y
1486	Publ Res Inst Math Sci	Project Euclid	y
1363	Sugaku		y
1242	Hiroshima Math J	Project Euclid	y
1082	Hokkaido Math J		y
1028	Funkcial Ekvac	J-STAGE	y
972	Kodai Math J	Project Euclid	y
952	Tokyo J Math	Project Euclid	y

Table: Journal titles which have over 150 articles.

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	ProjectEuclid	y
315	Kyushu J Math	J-STAGE/IR	y
297	J Math Sci Univ Tokyo	IR	y
271	Sci Rep Tokyo Kyoiku Daigaku Sect A		n

Concept of Digital Mathematics Library

- 全ての成果を電子的にアクセス可能に！ (2000 頃) All mathematical articles should have electronic version.
- コミュニティの役割が期待された。Community based digitization was expected.
- しかし、(特に日本では) 多様性が仇になる。Diversity includes difficulty.
- デジタルリポジトリの役割。Role of digital repositories.

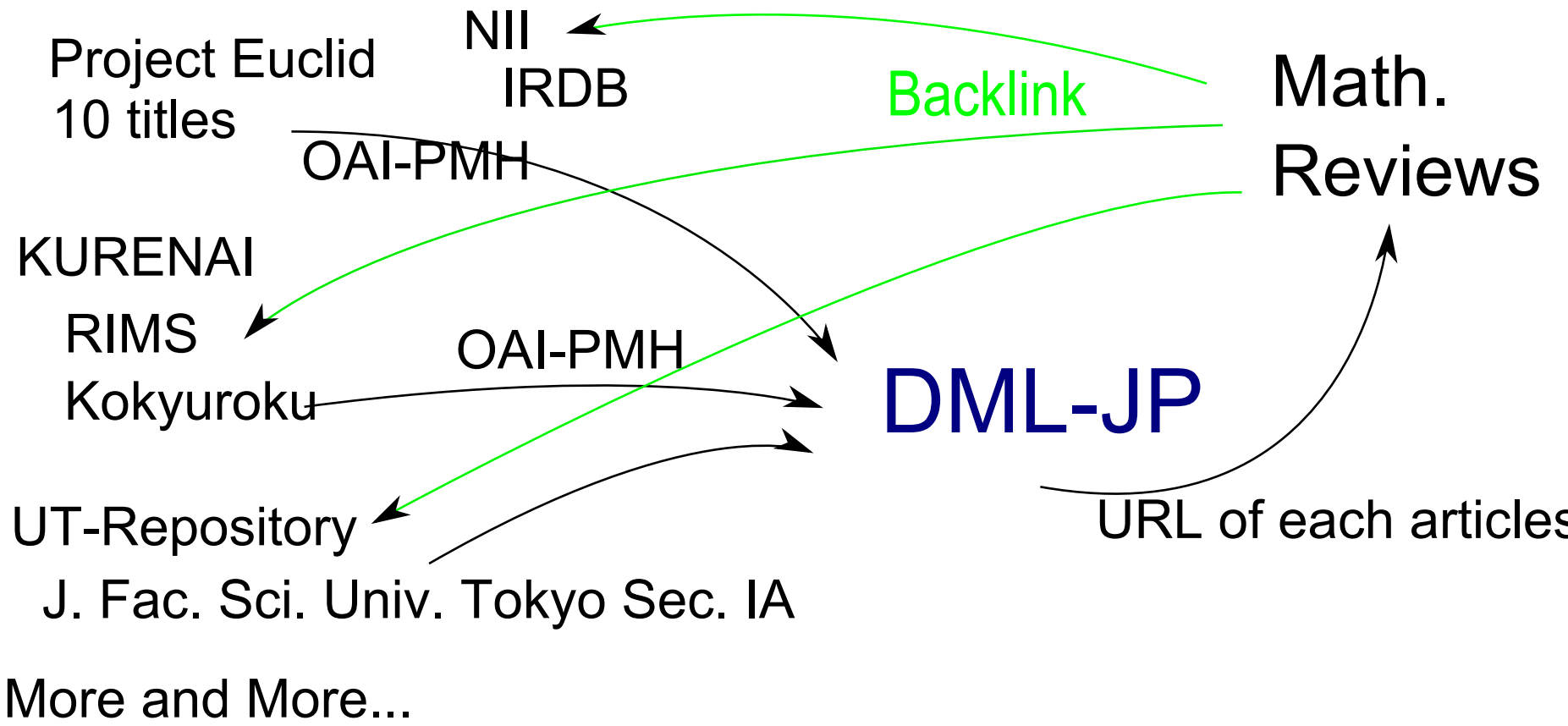
Digital Mathematics Library in the world

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)

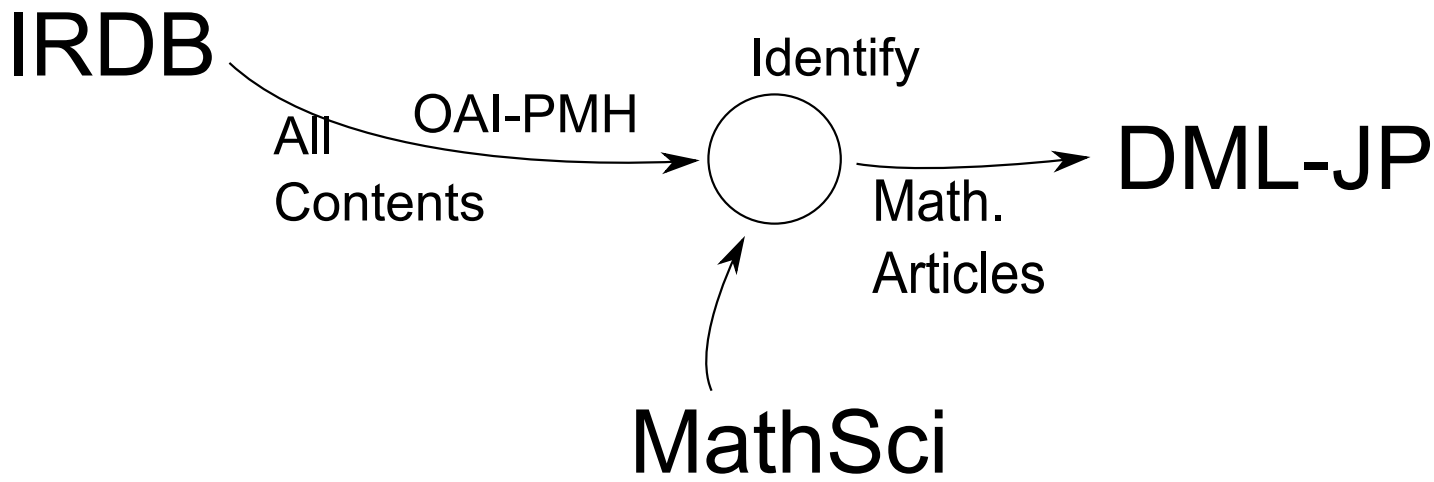
DML-JP, Digital Mathematics Library, Japanese part

- 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.



紀要も 重要/KIYOU is important

- For interdisciplinary titles, especially KIYOU, matching with Math Reviews database



Backlink from Math. Reviews/MathSciNet to original repositories can be made.

プレプリントの役割/Role of preprint

- 証明をチェックすると査読に時間がかかる。Review process take several month or some years in mathematics.
- 掲載誌発行以前のコミュニケーション手段が必要。This means that it takes certaian period to publish an article after submission even though electronic era.
- 数学教室による「プレプリントシリーズ」の発行。So mathematics department publishes "preprint series" for early communication and priority by exchanging.
- Many articles have been cited before publishing as preprint, like such a form "*author*, *title*, **preprint**". There are difficulties to find published version like

  [2] P. Alsholm, Wave operators for long-range scattering, (preprint, 1976).

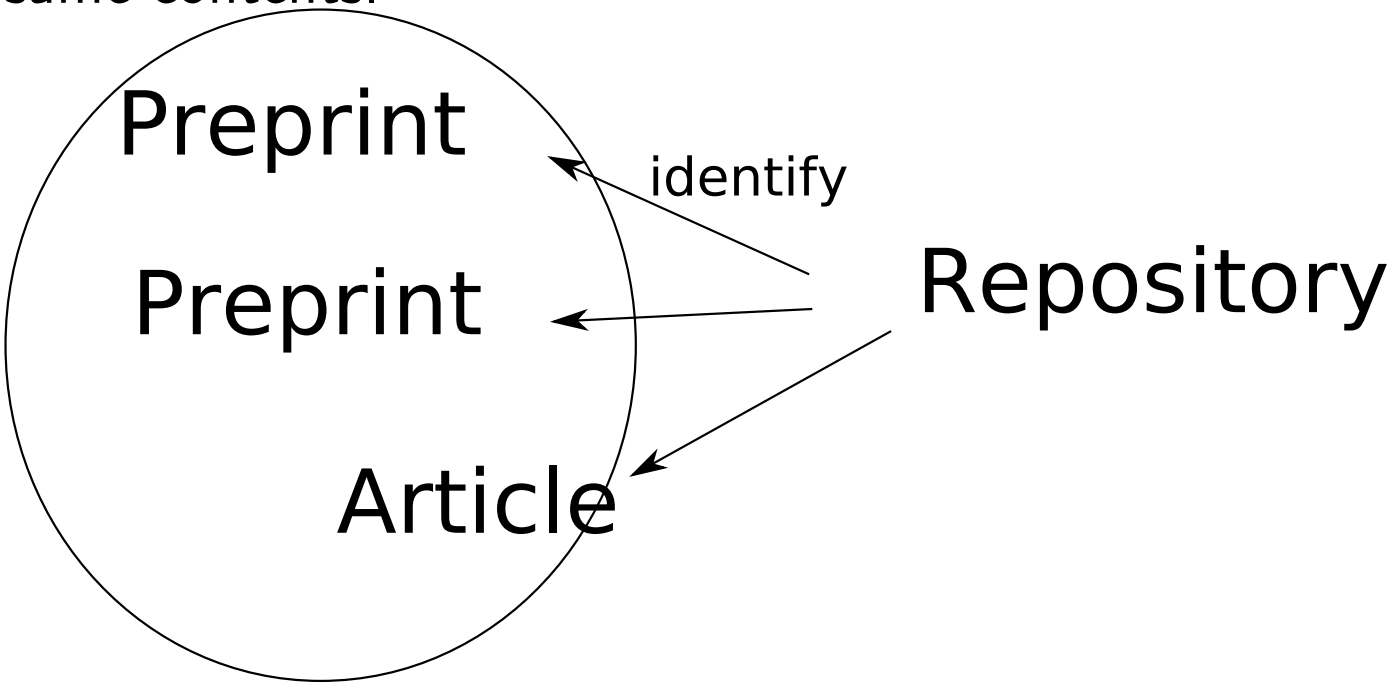
- By mathscinet search, we have "J. Math. Anal. Appl. 59 (1977), no. 3, 550–572".

Preprint, modern style

- arXiv.org の役割は伝統的なプレプリント 交換の効率化。The role of arXiv.org, what they say central subject repository, is to refine traditional communication style in digital era.
- Many people says arXive.org is a model of Open Access, however, actual recognition is not.
- 「プレプリントシリーズ」のオーバーレイ現象。Like `http://arxiv.org/abs/0812.3614`, part of electronic version of departmental preprint series now on arXiv.org using *report-no* field.
- Through OAI-PMH interface we can get

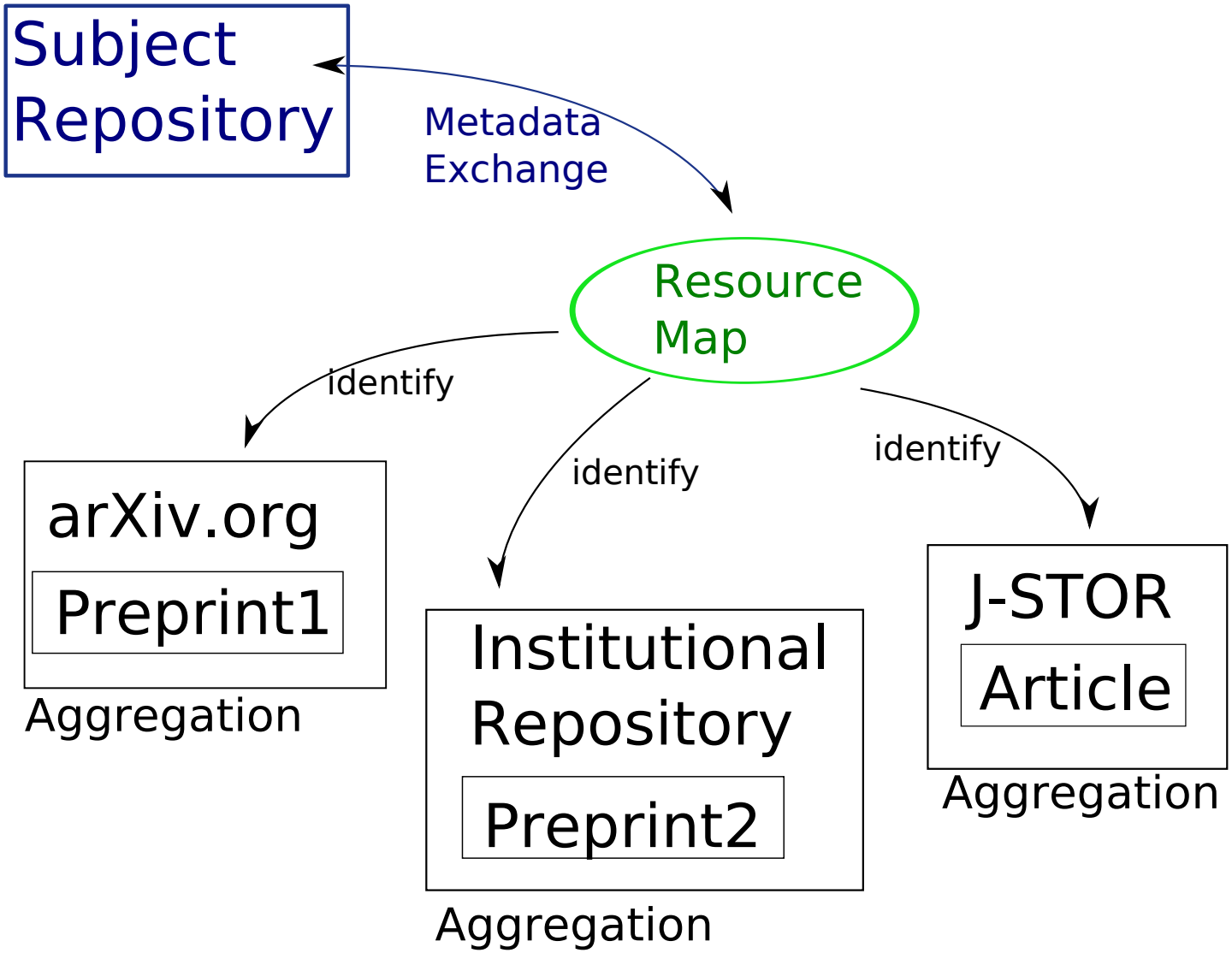
```
http://export.arxiv.org/oai2?
  verb=GetRecord\&metadataPrefix=arXiv
  \&identifier=oai:arXiv.org:0812.3614
```

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.



OAI-ORE realization

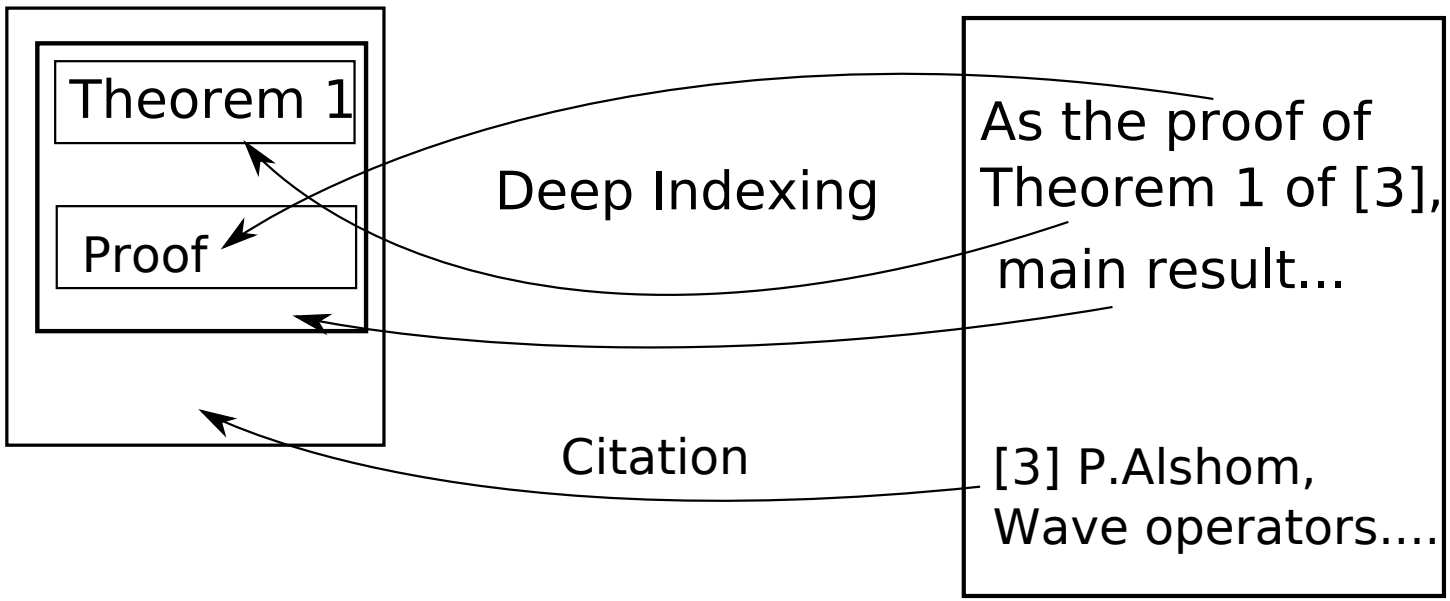
同定できればOAI-OREで記述できる。Once identified, we can describe by OAI-ORE.



将来的に何が起きそうか。In the future ?

- In the context of Open Access Movement, we need that "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)
- For example, citation is found like `[Theorem 1, Fermat's last theorerm]`, so we need deep indexing for theorems, proofs and so on.

Deep indexing diagram



An realization

```
<theorem no="1">
<statement>
For all natural numbers n>2 there does not exist x,y,z
which satisfy the following equation:
<math>
  <msup><mi>x</mi><mi>n</mi></msup><mo>+</mo>
  <msup><mi>y</mi><mi>n</mi></msup><mo>=</mo>
  <msup><mi>z</mi><mi>n</mi></msup>
</math>
</statement>
<proof>
  We have no space to describe the proof...
</proof>
</theorem>
```

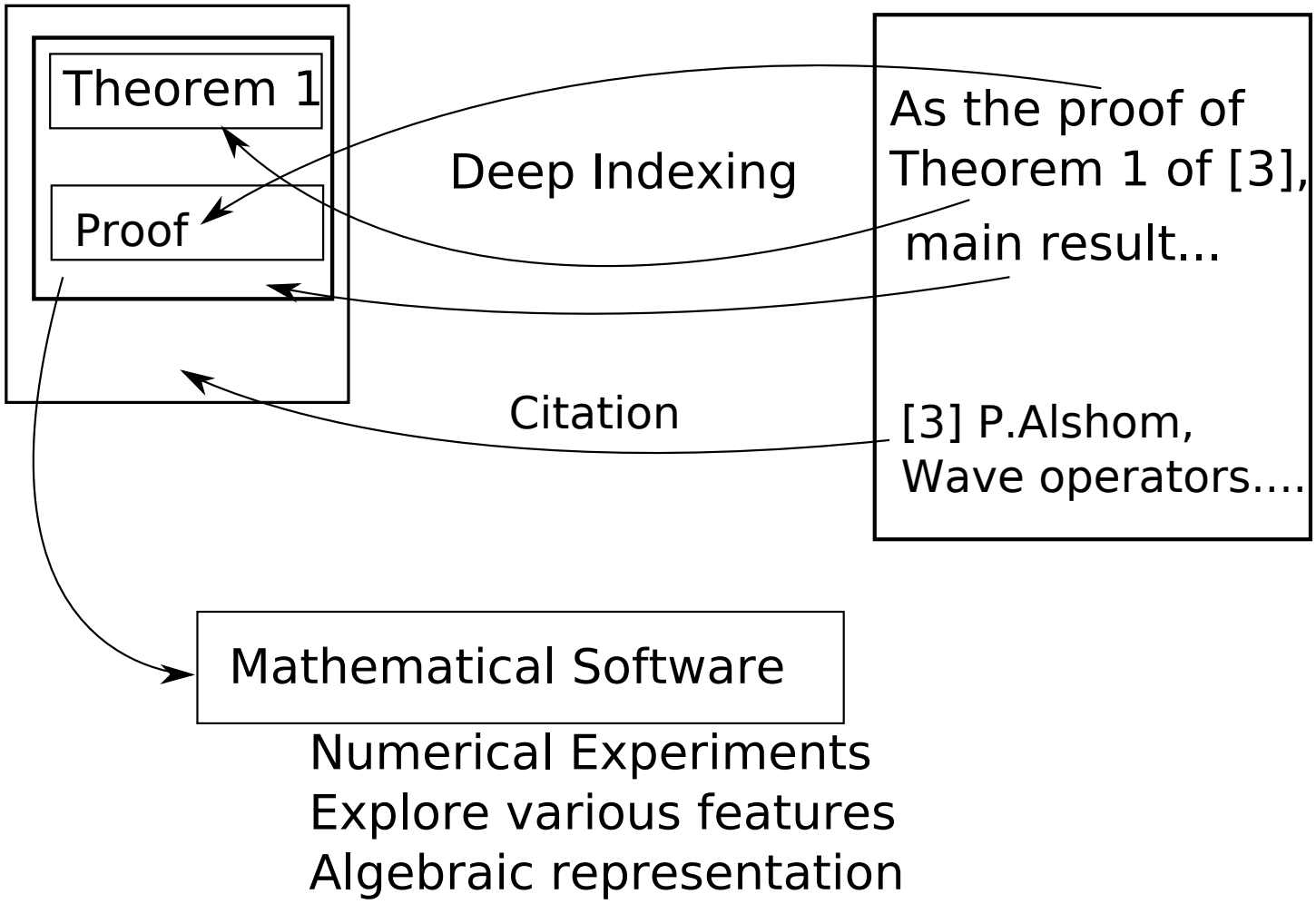
What is eScience in Math. ?

- Numerical experiments.
- Mathematical Software.
- Examples for unsolved problems.
- ...

A role of Math. in eScience

- Mathematical analysis on applied situation.
- Example: Computational Homology
- http://chomp.rutgers.edu/database/

Paper to Software



Conclusion

- 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。More collaboration !