

Lower Saxony State and
University Library Göttingen



Searching for Mathematics

Layers, Facettes, and Dimensions

Dr. Thomas Fischer

Metadata und Databases

SUB Göttingen

fischer@mail.sub.uni-goettingen.de

IMA Workshop

Enhancing the Search in Mathematics

Minneapolis, 26/27-04-2004

Overview

- ▶ This is a vast and complex area.
- ▶ There are a lot of activities going on.
- ▶ I can only give a indication of the things I am aware of, many other things might be known in the audience. Contributions are very welcome!

Search in general: multi-layered

Searching in a scientific context runs through several stages or layers:

1. Information is needed for a particular question; in a specific context,
2. Find relevant resources
 - ▶ use references in related material
 - ▶ use classification, keywords
 - ▶ text-based full-text search
 - ▶ not possible yet: search for formulas, patterns etc. (but probably e.g. for $\text{T}_\text{E}\text{X}$ expressions)

General Search (continued)

3. Find full reference for resources (where is it?)
(preferably optimized for local situation)
4. obtain referred resource (get it)
(in electronic or printed form, choice of digital format?)
5. extract information (read it)

Conceivable Vision: Semantic Shortcut

Shortcut from 1 to 5:

- ▶ Knowledge Management System represents the given knowledge in an appropriate database
- ▶ Definitions, theorems and proofs are represented semantically to obtain information immediately (RDF?)
- ▶ Relevant references can be given, but are not necessary.

This provides required information directly as mathematical knowledge.

→ Larger vision: Semantic Web

Facets: Special Features of Mathematics

1. Highly developed infrastructure:
Reference Journals, Preprint Servers, Math-Net
→ specialized search engines: BASE Mathematics, MPRESS, Google with DSpace
2. Formulas, symbols: $\wp \otimes$, Specialized Formalism
→ Formula recognition, e.g. INFTY
3. Highly internationalized
→ International standards, e.g. Name Authority Files (LoC, DDB)
4. High importance of legacy material
→ Digitization efforts: WDML, JSTOR, EMANI, NUMDAM, Elsevier, Springer...

Dimensions: Different Material

1. Printed Material

- Organization, distributed collections? Price of journals...
- Distribution as Xerox copy or scanned images (ARIEL)
 - fast delivery services, e.g. GAUSS
(http://www.sub.uni-goettingen.de/ebene_2/2_fl-pre.htm)
 - intellectual property rights (Germany/USA...)

Dimensions: Different Material (2)

2. Retro digitized books and articles

- Keep track of digitization efforts
 - Digitization Registry
(e.g. <http://www.sub.uni-goettingen.de/ssgfi/digreg/>)
- Create unified access for search and retrieval
 - Metadata standards, Application profiles
- Enhance search options
 - text recognition for mathematics

Dimensions: Different Material (3)

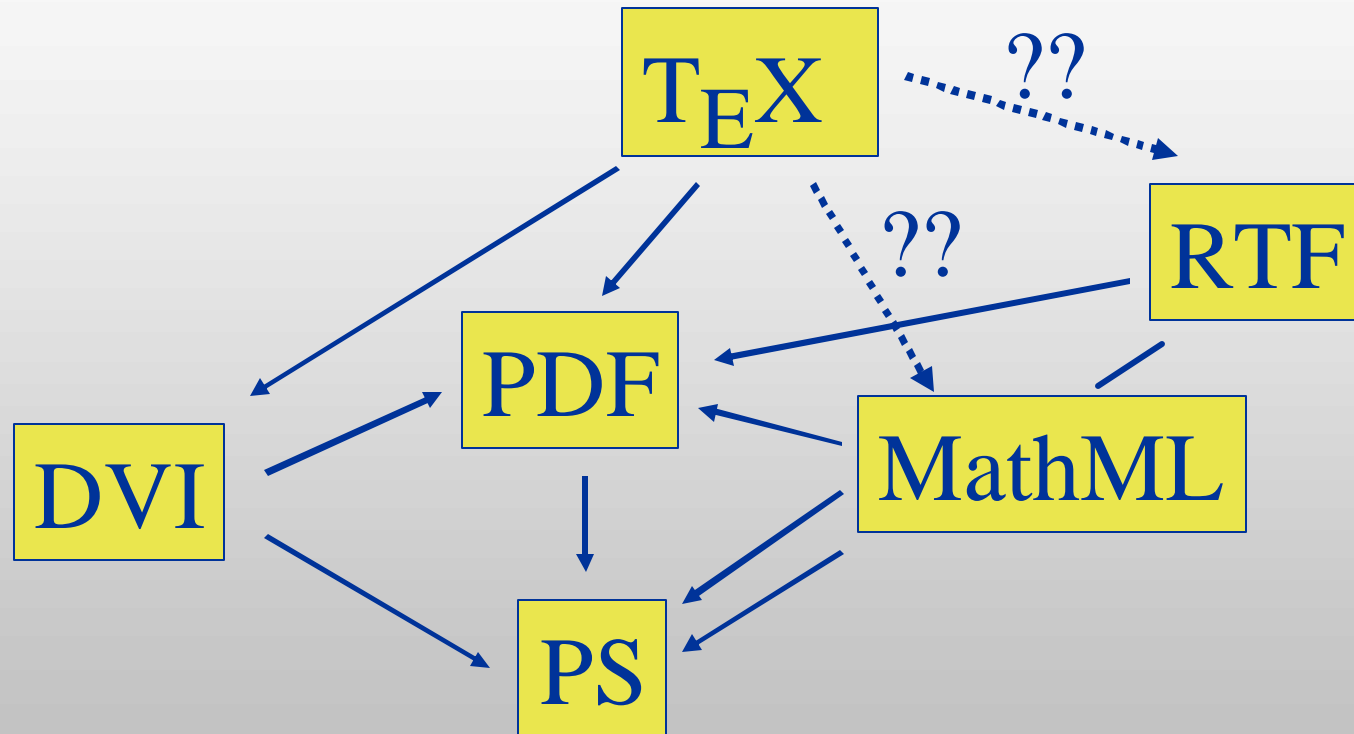
3. Born digital material

- Preprints, eJournals, eBooks:
access, ownership, reuse
→ SPARC, Budapest Open Access Initiative
- Different formats: TEX, MathML, DVI, PostScript, PDF, OpenMath, Mathematica, Maple ...
→ Conversion tools, semantic enrichment
- Archiving questions: long term availability for all or some formats (emulation?), trusted repositories
→ Conversion tools, standardization, OAIS

4. Data sets and similar material

???

Text Formats in Mathematics



Not a commutative diagramm!

Some Examples

Books

- ▶ "Steenrod's Reviews" (Reviews of papers in algebraic and differential topology, topological groups and homological algebra / classified by N. E. Steenrod)

Reference Journals

- ▶ MathReviews (<http://www.ams.org/mathscinet/>)
(MathSciNet: Mathematical Reviews on the Web)
- ▶ Zentralblatt MATH (<http://www.emis.de/ZMATH/>)
- ▶ The Jahrbuch Project: Electronic Research Archive for Mathematics (ERAM)
(<http://www.emis.de/projects/JFM/JFM.html>)

Preprint Servers

- ▶ ArXiv Math (<http://front.math.ucdavis.edu/>) or arXiv.org Mathematics e-Print archive (<http://arxiv.org/form/math?MULTI=form+interface>)
- ▶ Mathematics Preprint Search System (MPRESS,
<http://MathNet.preprints.org/>)

Some Examples (2)

Subject Gateways

- ▶ EEVL - MathGate (<http://www.eevl.ac.uk/mathematics/>)
- ▶ MathGuide (<http://www.MathGuide.de>)
- ▶ BUBL LINK / 5:15 (<http://www.bubl.ac.uk/link/mat.html>)

Search Engines

- ▶ CiteSeer (Scientific Literature Digital Library,
<http://citeseer.ist.psu.edu/>)
- ▶ BASE Mathematics (<http://math.ub.uni-bielefeld.de/>, not yet publicly available, see e.g.
<http://www.diglib.org/forums/Spring2004/lossau0404.htm>)
- ▶ Google with DSpace (to come, see e.g.
<http://chronicle.com/free/2004/04/2004040901n.htm>)

So much for now...

Questions?

Dr. Thomas Fischer
Metadata and Databases
SUB Göttingen
fischer@mail.sub.uni-goettingen.de