

ERAM, EMANI, etc.



Digitization projects in mathematics

Bernd Wegner

Mathematics Institute, TU Berlin

Editor-in-Chief of Zentralblatt MATH

Co-ordinator of EMANI, EMIS and ERAM

*Electronic Mathematics Archiving Network
Initiative*

Overview



- 1. ERAM
- 2. EMANI
- 3. RusDML
- 4. Additional projects

ERAM (Jahrbuch Project)



- Funded by DFG
- Electronic Research Archive in Mathematics
- Cooperation between SUB Göttingen and TU Berlin
- Capture the content of the Jahrbuch über die Fortschritte der Mathematik (1868 - 1943) in a database
- Digitise journals from the Jahrbuch period
- Technical infrastructure at GDZ
- 1,2 Million pages digitised

ERAM selected publ.



- Aequationes mathematicae
- Commentarii mathematici Helvetici
- Geometric and functional analysis
- Inventiones mathematicae
- Mathematische Annalen
- Mathematische Zeitschrift
- Journal für die Reine und Angewandte Mathematik (Crelle`s Journal)
- Metrika
- Numerische Mathematik
- Semigroup forum

ERAM selected publ.



- Monatshefte der Mathematik
- JACO
- EMIS backvolumes:
 - Czech J. Math
 - Publ. Math. Inst. Belgrade
 - Math. Scand.
 - BAG and others
- German dissertations
- Klein`s Encyclopaedia of Mathematics

EMANI - Aims & Scope



- Storing existing digitized mathematical content
- Digitizing older material
- Defining functional and user requirements
- Developing models and methods to achieve long term preservation
- Identifying tools for archiving, accessing, and searching
- Recommending formats for metadata, full text, supplementary material
- Securing rights
- Developing awareness
- Creating acceptance

Current partners



- Tsinghua University Library
 - Network centre for electronic library services in China
 - Mirror site of LINK, EMIS, Zentralblatt MATH
 - Ongoing digitisation activities
- Cornell University Library
 - Centre of excellence for electronic services (Euclid)
 - Mirror site of EMIS and Zentralblatt MATH
 - Ongoing digitisation activities
 - Host of the ArXive, planning group of DML

Current partners



- SUB Göttingen
 - Comprehensive reference library for mathematics
 - Mirror site of EMIS, Zentralblatt MATH
 - Ongoing digitisation activities, digitisation centre
 - Partner in several projects on STI
- MathDoc: Orsay Mathematical Library/Cellule MathDoc
 - Coordinator of French Math.Lib. Consortium
 - Partner in several projects on STI
 - Ongoing digitisation activities (NUMDAM)

Action items



- Check the reliability and usability of the current digital content
- Define the most reliable sources for long-term preservation of electronic versions (TEX, PDF)
- Define a minimal version of the content and structure of a document to be preserved
- Develop extended metadata schemes which are appropriate to capture information about archiving matters

Action items



- Set standards for retro-digitised offers on different levels
- High-resolution images
- Reference links
- internal links
- Text OCR
- formula OCR
- Presentation structure mark-up
- Define schemes for the preservation of editorial and other general information related to the retro-digitised publications

Action items



- Develop tools and structures for the content mark-up of scanned images (reference linking)
 - Project Michler/Schneider/Dennis
 - Tool from Tim Dokchitser
- Develop ideas about the conversion of the content into alternative formats more convenient for preservation (MathML etc.) or display.
 - HERMES, tool from Romeo Angelache
- Quick access through reviewing services

Key problems



- Copyright: Get the approval from authors and publishers to provide access
- Availability: Develop an integrated access structure
- Extension: Develop a structure which is capable to integrate further partners
- Business models.
- Improvements: Conversions
 - Image to Latex
 - Latex to MathML

First results



- Extension of digitised content:
 - 1.200.000 pages in ERAM at GDZ
 - Communications in Math.Phys. at Cornell
 - Ongoing work at MathDoc
 - Chinese digitisation project
- TeXDocC software at SUB for TeX evaluation
- Seamless interlinking of digitised and digitally born content

EMANI content



- State and University Library Göttingen
 - 27 journals in ERAM
- Cornell University Library, Ithaca
 - 19 journals in Project Euclid
- Tsinghua University Library, Beijing
 - 55 journals in Chinese Math
- Cellule MathDoc, Grenoble
 - 4 journals in NUMDAM

МОСКВА-БЕРЛИН MOSKAU-BERLIN

государственный Исторический музей ИСКУССТВО ЗА ПОЛВЕКА

950
000

с 3 апреля
по 14 июня

МОСКАУБЕРЛИН
МОСКВАБЕРЛИН

RusDML Objectives



- Make all Russian mathematics electronically available
- Make all Russian mathematics easily accessible
- Disseminate information on Russian mathematics world wide
- Establish the Russian part of the DML or DML-EU – global distributed virtual libraries in mathematics
- Free access as far as possible

RusDML Scope



- All mathematical publications from Russia and FSU countries
- Take a set of about 70 journals for the core archive
- Later extension to monographs, collections of articles and deposited work
- Care about the “printed only period” at first
- Later extension to digitally born material

Russian Team



- GPNTB Moscow, main RusDML backbone on the Russian side
- Associated partners
 - RAS,
 - Steklov Institute of Mathematics
 - Supporting libraries,
- Content providers: Russian editors and publishers
- Application for funding at RFBR



German Team

- SUB Göttingen, RusDML backbone for digitisation and metadata
- TIB Hannover, RusDML backbone for digitisation and metadata
- TU Berlin / Zentralblatt MATH: provision of raw metadata, database links
- DFG – German Research Foundation: financial support

RusDML Workflow



- Preparation of metadata by Zentralblatt
 - Collect the information journal by journal
 - Add missing information
 - Add missing reviews
 - Correct errors in metadata
 - Standardise bibliographic information
- Delivery of metadata to digitisation centres,
 - reviews included,
- Digitisation at SUB and GPNTB
 - Immediate generation of permanent URLs



RusDML Workflow

- Preparation of final metadata in Russian and English
 - Structural data
 - Archiving data
 - Technical descriptions
- Installation of bilingual access
- Distribution of content to all partners (and additional parties of interest)
- Parallel and preparatory activities
 - Development of tools and structures
 - Caring about rights and licenses

Digitization activities



- NUMDAM (NUMérisation de Documents Anciens Mathématiques), Cellule MathDoc, Grenoble
- DIEPER (Digitised European Periodicals), European project, - Mat. Sbornik
- Gallica, collection of retrodigitized material, BNF
- JSTOR, Elsevier Backfiles, Springer SBM Project
- ERAM
- Cornell University Math Collection
- Digitised Chinese Mathematics, Tsinghua
- 16 Korean math journals at KISTI

Digitization projects



- RusDML
- DML-CZ
- Polish DML
- Catalan DML
- Planned or starting projects in
 - Spain, Portugal, Italy, Switzerland
 - Greece, Bulgaria
 - Australia? Great Britain? Nordic countries?
 - India ??
- Still open: Hungary, Romania

Adresses



- Thank you for the attention
- <http://www.emani.org>
- E-Mail: Wegner@math.tu-berlin.de
- <http://www.emis.de>