Ongoing Work on a
European Virtual Library of Mathematics
General presentation

Thierry Bouche

Cellule MathDoc & institut Fourier,
Université de Grenoble
Electronic Publication Committee of the
European Mathematical Society

Towards a Digital Mathematics Library
Grand Bend, Ontario, 8-9 July 2009
Math literature

The mathematical literature

Stakes

- Mathematical *validated* literature never becomes obsolete
- Old results are not superseded by newer ones: they are their foundation
- It’s valid only as a *whole*, building a wide network of references
- It’s useful to other sciences in an *asynchronous* fashion
- It must be carefully archived, indexed and preserved
- It must be accessible over the long term
We thus need a reference library, which should be

- exhaustive
- up-to-date
- well organized
- widely open
- easy to use for non-mathematicians

**Paper** OK? (libraries, ILL, doc. delivery, union catalogs, reviewing DB. . . )

**Electronic** Still a dream! (WDML)
The mathematical literature

Time scale

- Instant preprint circulation (labs, arXiv, email, home pages)
- Actual publication delayed 1-2 years
- Publication’s goals: prestige, attribution, quality rating...
  To secure the version of the work suitable for further reference
- About 50% of citations in today’s bibliographies are more than 10 years old
- About 25% of citations in today’s bibliographies are more than 20 years old
The mathematical literature

Size

A rough estimate on the size of the whole corpus of written mathematics in the occidental scientific tradition since Euclid:

- 3 million items were published spanning < 100 million pages
- 100,000 new items appear each year
- < 20% published before 1900
- > 50% published after 1950
- 80% journal articles, 10% chapters in collective books, 10% books
- 600 math-only journals alive, 2000 with math articles
- 15 million pages digitised? 65% of core journals available digitally?
The mathematical E-literature

Needs

- Going electronic *should* be a wonderful asset for opening new ways of using the mathematical corpus beyond old boundaries.
- The main infrastructure required would provide the basic features of the reference library, plus e-only add-ons
- This means
  - A global (distributed) facility dedicated to archive newly published or digitised material
  - An up-to-date registry of all available resources
  - Mechanisms for interlinking the holdings with existing and future infrastructures
  - Seamless navigation across the whole corpus
Electronic media has downsides for scholars and librarians

- Costs increase!
- Many new access barriers (copyright, licences, DRMs)
- No standards for interfaces, file formats, etc.
- Mainstream publishing is not adapted to mathematical content. . .
- “Value” is measured by counts (not scientific value)
Many paper items are missing a digital counterpart, \textit{but}.

Many digital items are duplicated among various providers, \textit{while}.

Many collections are split across providers, \textit{and}.

Collection holders are very volatile.

\implies Managing an exhaustive and up-to-date access requires zillions of subscriptions, and superhuman monitoring capabilities.
A reference digital mathematics library should assemble as much as possible of the digital mathematical corpus in order to

- **preserve** it over the long term,
- **make it available online**
- **at reasonable cost,**
- in the form of an **authoritative and enduring** digital collection,
- **growing** continuously with publisher supplied new content,
- **augmented** with sophisticated search interfaces and interoperability services,
- **developed and curated by a network of institutions**

**Challenge:** define the highlighted terms in such a way that a sufficient diversity of stakeholders representing a critical mass commit themselves to the envisioned effort.
Previous work on coordination


Digital Mathematics Library. NSF planning project (2002-2003, Cornell University Library) “toward the establishment of a comprehensive, international, distributed collection of digital information and published knowledge in mathematics”.

Mathematical Knowledge Management meetings (2001–) + DML workshops (2008–): technical challenges.

EMS’ EoI to the European Commission (2003), supported pilot implementation proposals to EC programmes (2003–2009: FP6, eContentplus, CIP ICT PSP . . .)

AMS/MSRI proposal to the Moore foundation (2005)

The Virtual Library of Mathematics
Existing content

America  JSTOR (260,000 items), project Euclid (100,000), CMS (4,000)
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 items)
  Poland  ICM/BWM (13,000 items)
  Portugal  SPM/BNP (2,000 items)
  Spain  DML-E (5,000 items)
  Czech Rep.  DML-CZ (11,000 items)
  Russia  RusDML (13,000 items)
  Bulgaria  BulDML (2,500 items)
  Serbia  No formalised project (3,700 items)
  Switzerland  SwissDML (5,000 items)

Commercial  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)
The Virtual Library of Mathematics

What remains to be done?

- Give up the world-wide top-down approach to the effort
- Identify a core group of stakeholders with different backgrounds willing to go ahead
- Learn from success stories of existing projects and generalise them
- Call the scientific community to support and help shape the effort
- Define balanced, inclusive policies for an ever-growing sustainable infrastructure
- Get much better return on investment for Research organizations through European momentum
- Give a new impetus to research communities in Digital libraries and Mathematical knowledge management to design a much more powerful yet reliable research environment