World Digital Mathematics Library

IMU-Panel, Seoul ICM 2014, August 20

What do we need?
What can reference databases provide?

Gert-Martin Greuel

University of Kaiserslautern, Zentralblatt MATH (zbMATH)

General Remark

Mathematical knowledge

- Although a complex system, it is still primarily based on scientific publications (other sources such as mathematical databases and software become more important).
- Mathematical knowledge may be considered as a continually growing huge building in which it is necessary that each floor is reached and no stone is lost.
- In mathematics research, the literature plays a very special role, due to the timeless validity of its achievements that is mainly preserved in the published scientific literature.
- Since mathematical knowledge does not become obsolete, it is therefore necessary for the researching mathematicians as well as for the users of mathematical methods to have access to the entire mathematical literature.
- The published mathematical literature is the main starting point for a WDML but it should be open to other emerging aspects of mathematical knowledge.

In the following I talk mainly about the **mathematical literature.**

What do we need for a WDML?

1. digitize mathematical literature that is not already in digital form
(a big amount has been digitized by a variety of different commercial and non-commercial providers)

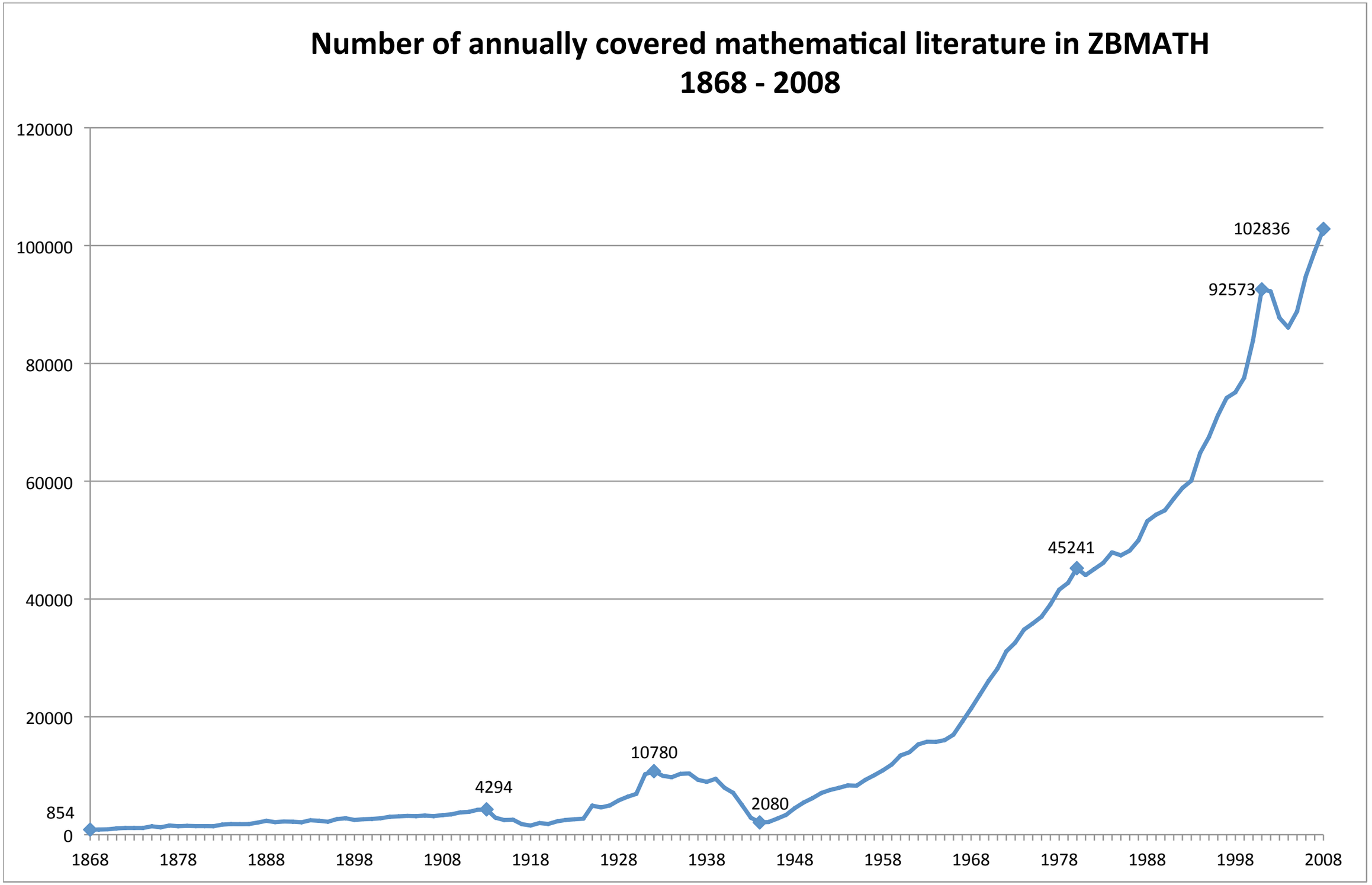
2. technical standards for making digital mathematical literature accessible online
(there exist different formats for data and metadata of a large number of digital libraries organized by different providers under different access conditions)

3. architecture for making digital mathematical literature and mathematical knowledge in general available in the future (formats, metadata, and knowledge management in general are under permanent development)

Major problems

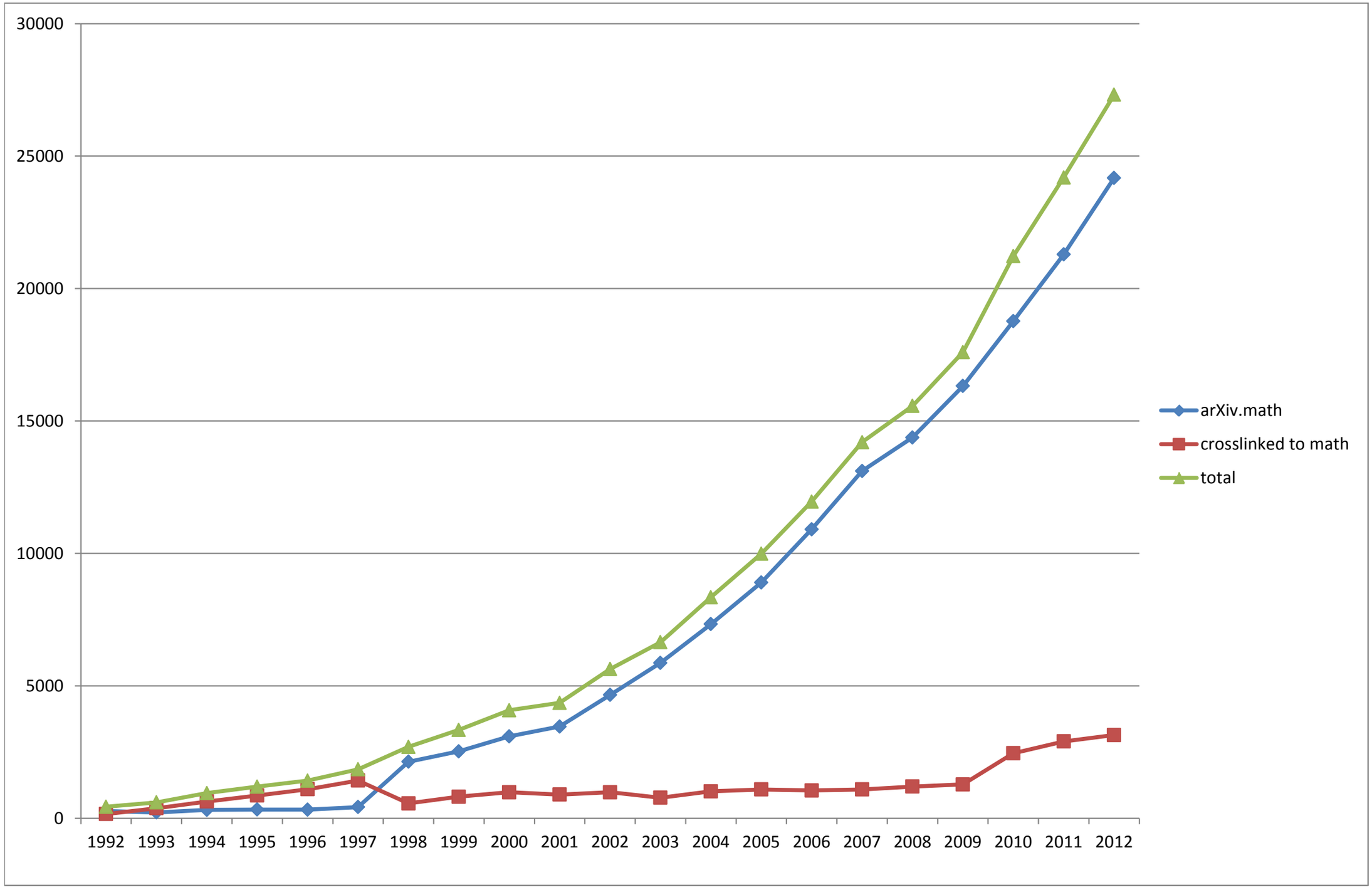
1. **Content** (deciding what has to be included and what not)
2. **Initial Format** (to enable global features for presentation, searching, computing)
3. **Copyright** (clearing complicated legal issues in international copyright)
4. **Archiving** (technical format for archiving)
5. **Long term funding** (business model)

A challenge: the growth of the mathematical literature



Growth of (peer reviewed) mathematical literature is a challenge also for the reviewing services. (The decrease between 1938 and 1948 is due to WW II, the one between 1998 and 2008 has mainly internal reasons. In 2013 the number went up to about 120.000 items). There is no indication of a decrease of the growth rate.

Growth of preprints in ArXiv.math



The arXiv is growing even faster, mainly because it is much newer; cross linked papers are primarily published in other areas than mathematics (data are taken from arXiv). arXiv: from 2007-2012 (5 years) the number has almost doubled, zbMATH: from 1988-2008 (20 years) the number has doubled.

Reference databases as potential access infrastructure to the WDML

Recall the IMU/CEIC requirement: Each article should include a separate list of references with links to the indexing databases Mathematical Reviews and Zentralblatt Math

Reference databases have several advantages:

- Provide identifiers for the indexed mathematical literature
- completeness of mathematical literature
- high quality and well-structured metadata, math specific search options
- exclusive to mathematical literature, little noise
- semantic content analysis (MSC, keywords, abstract, reviews)
- linking of information (e.g., with full texts, if available), references
- author disambiguation, author profiles

Reference databases are engaged in the development of tools for the WDML

- decide continuosl what to index (i.e. what is math literature)
- development of metadata schemes for mathematical publications
- maintenance of the Mathematical Subject Classification (MSC)
- pilot partner for the use of the methods for publishing and presenting mathematical knowledge (e.g., use of MathML as presentation format)

Hence reference data bases can provide core services for the WDML