

Math-Net

The State of the Art of a Distributed Information & Communication System in Mathematics

Winfried Neun

Konrad Zuse Institute Berlin (ZIB), Berlin, Germany
neun@zib.de

Abstract. The Math-Net Initiative (<http://www.math-net.org>) steered by the International Mathematical Union (IMU) tries to improve and to coordinate the mathematical scholarly information in the World Wide Web. Math-Net is a community-based information and communication system in mathematics. It is based on the information which is provided by persons and institutions taking part in Math-Net on their local servers (Math-Net Members). They should make their information resources electronically available in a standardised fashion. Currently the Math-Net activities are focused on three topics: Math-Net Pages, Mathematical Preprints, and Personal Information about Mathematicians. The talk will give an overview about the idea of Math-Net.

The Math-Net Page, a standardised portal to the information of mathematical institutions, will be explained in more detail.

1 Introduction

The origin of the Math-Net lies in developments of the last decade, namely the enormous increase of information made available on the Net on the one hand and a significant increase of the prizes for traditional mathematical communication means like journals or textbooks on the other hand. This made the usage of the World Wide Web for the exchange between researchers very plausible. For a scientific community, the Web contains a huge number of relevant information provided by many individuals and institutions. But often not all of the relevant information can be found in the Web quickly and it may be difficult to obtain the required information at all. The Math-Net Initiative (<http://www.math-net.org>) tries to improve this situation in mathematics.

2 The Math-Net concept

2.1 The Aim

The general aim of the Math-Net activities is to improve the access to the mathematical information for the user.

Math-Net is a community-based information and communication (short *i&c*) system in mathematics. It is a not-for-profit activity steered by the International Mathematical Union (IMU), the world association of mathematicians. The aims, the principles, and the organisation of this initiative are defined in the Math-Net Charter, which can be found e.g. in [1].

From a technical point of view, Math-Net is a structured, distributed, interoperable, user-friendly, and high quality electronic information and communication system.

Any institution, person, or group of persons willing to make its mathematical resources electronically available within the scope of Math-Net may become a Math-Net Member. Math-Net bases on the information which is provided the Math-Net Members on their local servers. They should make their information resources electronically available in a standardised fashion. Math-Net Members have full responsibility for the quality, accuracy, timeliness, and appropriateness of the data they contribute.

Math-Net services combine the data made available by the Math-Net Members into services. These services aim at providing fast and well-structured access to the mathematical resources within Math-Net (and possibly beyond).

In detail, the Math-Net Initiative engages itself to

- structuring, organising, and standardising the information offered by participating persons and institutions (as technical base XML and RDF are used)
- describing and analysing the contents of objects and links,
- indexing data and metadata,
- archiving material of long-term interest,
- developing and applying high-quality presentation and authoring methods and tools,
- developing enhanced methods for retrieval
- insuring software interpretability and interdisciplinary compatibility,
- improving scientific information services.

The Math-Net Initiative stems from the Math-Net project (1997 - 1999) in Germany. Leading this project to an international level is one important goal that is rather challenging. Currently the Math-Net activities are focused on three topics:

- (institutionnel) Math-Net Pages,
- preprints, and
- personal information about mathematicians.

2.2 Information provided by the Math-Net Members

Math-Net is a distributed information and communication system: the information offered by the Math-Net Members on their own Websites is the base and the core of Math-Net. Math-Net Recommendations and Standards define guidelines for a standardised editing and description of the digital information.

2.3 Math-Net Services

Math-Net Services are general, type- or topic-specific portals to the information of the Math-Net Members. Math-Net Services gather the local information, and make it accessible in a user-friendly form. The current Math-Net Services are: SIGMA, NAVIGATOR, MPRESS, PERSONA MATHEMATICA, MATHJournals and MathLinks.

2.4 Organisation

The Committee on Electronic Information and Communication (CEIC) established by the IMU in 1998 coordinates all Math-Net activities. The CEIC coordinates the activities of the mathematical community in the field of information and communication. The Technical Advisory Board (TAB) is responsible for the technical progress in Math-Net. Each Math-Net Member, represented by its Information Coordinator, participates in the advances of Math-Net through Math-Net Member Associations.

3 Technical aspects

3.1 Metadata

The enormous growth of information in the Web makes it necessary to develop new techniques for publishing and managing the information allowing a processing of the information by machines. The Semantic Web is the idea for a machine-understandable processing of information. Metadata play a key role here. Metadata, "data about data", provide additional information for a special purpose. One example for a metadata standard is the Dublin Core metadata: In the mid-nineties librarians started to develop a set of bibliographic data for a standardised description of digital documents-like objects. The result was the Dublin Core metadata set containing 15 elements to characterise digital publications.

Metadata allow a more precise search within the documents compared with a fulltext search.

In a HTML encoded document the Meta-Tag was used to carry this information. But HTML is too restrictive for metadata: it does not allow to develop flexible metadata models oriented at the needs of the users.

The development of the Resource Description Framework (RDF) is a milestone for the Semantic Web and a universal concept for metadata. At first, RDF is a graph model for the definition of metadata. Moreover, RDF/XML is the syntax to transport the model. RDF uses XML for the encoding of the statements. Moreover, RDF Schema provides a lot of additional functionalities for RDF/XML metadata. Especially RDF Schema allows the validation of statements or the definition of hierarchies. RDF Schemas can be used also for the definition of ontologies.

Math-Net uses XML and RDF in the following sense: Math-Net analyses the content of documents containing important types of information, especially the Math-Net Pages, preprints, and personal homepages. Math-Net Pages, preprints, and personal homepages are RDF classes. This means, a RDF metadata model was defined for each of these document types. Therefore we used as far as possible, metadata elements developed by international standards such as the Dublin Core Metadata Set for documents (publications) or the vCard Standard for persons. Further metadata elements were defined if necessary. Therefore, own namespaces were defined (URI, which allow the parsing of metadata). The metadata (and partially the documents) were encoded in XML. The whole XML metadata code for a specific application defines its application profile.

Math-Net Services harvest the information available through the Math-Net Pages, especially the metadata, and use the metadata to provide efficient access to the information. A metadata schema must be developed for each application. Namespaces allow the use of various tags coming from different namespaces.

The following sections provide some scenarios how metadata are used in Math-Net.

3.2 Math-Net Pages, the Math-Net Page Maker, and the Navigator

The idea of the Math-Net Page. Today, nearly all mathematical institutions world-wide provide information about themselves on Web sites. The Math-Net Page is an easy and straightforward idea for a uniform access to the information at the Web sites of mathematical institutions: It is a standardised (additional) homepage ('a portal') for the Web sites of mathematical institutions. The standardisation refers to

- the structure,
- the vocabulary,
- and the layout

of the Math-Net Page.

In a first step, the information on the Web sites of mathematical institutions (mathematical departments, research institutes and societies) was analysed and classified. The information was divided into 6 groups and subgroups and standard labels were assigned to these groups. Moreover, the Math-Net Page provides easy access to other local information such as the Web-sites of the University and especially to further Math-Net Services.

Mathematisches Institut
Universität Köln

Local Contact Local Search Homepage Univ. Homepage

General More Links

- ▶ About us
- ▶ Organization of the Department
- ▶ Information for Prospective Students
- ▶ Community Outreach
- ▶ Information for Visitors

People More Topics

- ▶ Faculty / Staff
- ▶ Students

News

- ▶ Schedule of Events
- ▶ Information for Dept. Members
- ▶ Positions Available

Research

- ▶ Research Groups
- ▶ Preprints / Publications

Teaching

- ▶ Academic Programs / Curricula
- ▶ Class Schedules
- ▶ Course Information and Materials

Information Services

- ▶ Computing Services
- ▶ Libraries
- ▶ Journals
- ▶ Bibliographic Databases

Standard Math-Net Page

© Mathematisches Institut, University of Cologne; last update: 2002-05-02

Fig. 1. The (standard) Math-Net Page of the university of Cologne

The basic idea of the Math-Net Page is a 'secondary homepage' for mathematical institutions, e.g. Math Departments, Research Institutes, or Mathematical Societies. This page allows the users to access the available information in a structured and even automated manner. The original ('primary') homepage of the institution, which is nicely designed in most cases, will be left as it is. On the secondary homepage the information of the institution is repeated and for the human reader it has an uniform appearance. All categories of the Math-Net Page are predefined in the recommendation. All these pages are created in the English language.

Internationalisation. In order to support the national mathematical communities, there is a need for Math-Net Pages in languages different from English. Btw: even among the institutions in English speaking countries there are slight differences in the wording of the categories. On the other hand, there are countries with

more than one official language, thus in some cases official Web Pages have to be drawn up in all official languages. To fulfil these requirements there is the concept of 'alternative' Math-Net Pages serving as a supplement to the standard page. These pages have the same design, but the institution is free to choose categories, language, etc. according to its needs. E.g. the German alternative Math-Net Page reflects some of the extra work accomplished during the first German project phase.

In general, Math-Net Pages, even standard pages in English, use multiple alphabets. Thus the encoding of the information has to be chosen wisely. The aim is, of course, to allow the browsers anywhere in the world to render the pages correctly. This may be trivial if one lives in the 'Latin-1' world, but in general it is not.

To solve such problems the W3C has recommended UNICODE [2], whose encodings are able to represent virtually all characters of all alphabets world-wide. The Math-Net pages allow for a usage of UTF-8 (one of the Unicode encodings) as an option. This is also a necessity for an automatic gathering of the information in the alternative pages.

Summing up, the internationalisation of Math-Net Pages means

- Defining one standard and alternative Math-Net Pages (the standard pages have common English names, Native language Math-Net Pages can be installed via alternative pages)
- Use of different character sets (especially UNICODE) The Math-Net Page provides different modes for handling different character sets:
 - The As-Is-mode: The character set the user will be applied.
 - The UTF-8 mode: The character set is changed to UTF-8, the web pages are stored using this encoding.
 - The default value for the character set is the As-Is mode.

Metadata. The Math-Net Page is a Math-Net resource class¹. A RDF model was developed for the Math-Net Page by IWI² and the Technical Advisory Board (TAB³) of the CEIC. It contains:

- metadata about the information coordinator
- metadata about the institution
- metadata relating to the groups and subgroups of the Math-Net Page.

The metadata defined by vCard Standard⁴ is used to describe the information coordinator and the institution. The metadata regarding groups and subgroups

¹ <http://www.iwi-iuk.org/material/RDF/1.1/Schema/Class>

² <http://www.imi-iuk.org/material>

³ <http://www.mathematik.uni-osnabrueck.de/TAB>

⁴ <http://www.w3.org/2001/vcard-rdf/3.0>

of the Math-Net Page are defined in the namespace: <http://www.iwi-iuk.org/material/RDF/1.1/descriptor/> The application profile for the Math-Net Page covers the XML representation of the metadata part at <http://www.iwi-iuk.org/material/RDF/1.1/profile/MNPage/>.

The Math-Net Page itself is an XML file. The visible part which is shown by the browser is coded in XHTML. The metadata of the Math-Net Page are a further part of the Math-Net Page.

The Math-Net Page Maker. Creating Math-Net Pages is a quite complicated and time-consuming task. Some reasons lie in the fact that it has to be guaranteed that the metadata information is present and that the information categories correspond to the recommendation. For the human user, the design should be almost identical, except for logos etc. It should also be rather simple to edit pages consistently, i.e. changing the visible information and metadata should effect at the same time.

Therefore, in order to facilitate and encourage the creation of Math-Net Pages, which at the same time improves the institutions acceptance in the community, a tools has to implemented. This tool which supports the local information coordinator in the institutions in creating the pages, consists mostly in filling in the information about the institution into a web based form. In most cases, the information is a set of URLs pointing into the pages of the local web server. This tool can be found at <http://elib.zib.de/cgi-bin/mpm.pl>.

The IMU Recommendation of the Math-Net Page. On 2002-04-12 the Executive Committee of the IMU and CEIC approved the recommendation on the Math-Net Page: "Based on the recommendation of its Committee on Electronic Information and Communication (CEIC), the International Mathematical Union (IMU) asks every mathematics institution throughout the world to create a Math-Net Page, to install a prominent link to that page from its primary homepage, and to maintain its Math-Net Page in the future."

The Math-Net Navigator. The Navigator is the Math-Net Service for the Math-Net Pages. It is a portal to the information provided by the Math-Net Pages. The Navigator collects the Math-Net Pages and indexes their metadata. The links to the local information are gathered and processed. As a result, the Math-Net Navigator generates different views to the Math-Net Pages e.g. about the institutions in a country or about the groups and subgroups of the Math-Net Page. See <http://www.math-net.org/navigator/>.

Additional remarks. The Math-Net Page plays an important role for the Math-Net Initiative. First, the Math-Net Page defines important topics for the develop-

ment of the Website offered by a Math-Net Member. Second, the Math-Net Page helps build up a personal infrastructure for the Math-Net Initiative. If a institution generates a Math-Net Page a information coordinator has to be appointed by the institution. She/he is the contact person for all Math-Net activities. At the beginning of the initiative, the Math-Net Page was designed for mathematical departments and institutes. Meanwhile a version for mathematical societies has been developed too.

Acceptance: More than 150 institutions world-wide generated and installed Math-Net Pages as of 2002-07-01.

3.3 Preprints and MPRESS

Preprints play a more and more important role in publishing scientific results in the last years. One of the reasons: Web and Internet allow a quick and effective way for a publication for the scientific community. As a consequence, a lot of mathematical institutions or subject-specific servers provide digital preprints in mathematics, all with their own rules for description. This makes it sometimes time-consuming to search a special preprint.

In the framework of the Math-Net project in Germany a metadata set for preprints basing on the Dublin Core metadata has been defined. Metadata of preprints should be generated as additional files, so-called shadow files. To generate the shadow files in an easy and effective way, the TU Osnabrück/IWI have developed the MMM tool (MyMetaMaker). See <http://www.mathematik.uni-osnabrueck.de/cgi-bin/MMM3.0.cgi>. The metadata part was revised in the last years: the HTRTML metadata was replaced by RDF metadata.

The University of Osnabrück/IWI provides also the Math-Net Service MPRESS, the Mathematics Preprint Search System: <http://mathnet.preprints.org/>. MPRESS is harvesting and processing the information about mathematical preprints from more than 120 Websites. In particular, MPRESS provides a browsing alongside the Mathematical Subject Classification. MPRESS is the most important portal for mathematical preprints. It contains information about more than 50,000 mathematical preprints, see: <http://mathnet.preprints.org/stats.html>.

3.4 Personal Math-Net Pages and PERSONA MATHEMATICA

Another important resource for the mathematical community is the information about people working in mathematics. Hence, the IMU has disseminated the printed World Directory of Mathematicians on the occasion of every International Congress of Mathematicians (this means every 4 years). The disadvantages are clear: The World Directory of Mathematicians is outdated at the time of publication. Changes

and Corrections are impossible. The personal information is restricted to a few topics.

Here the electronic media is a reasonable alternative. The idea is to make the World Directory of Mathematicians in the style of Math-Net. This means:

- developing a metadata model for the description of a person, possibly a personal Math-Net Page
- developing a tool for generating the metadata description by mathematicians
- local storage of the metadata/personal Math-Net Pages
- a search index (Math-Net Service) for the metadata of persons

Currently the concept for the World Directory of Mathematicians is under discussion.

3.5 MathML and OpenMath

In recent years two standards, MathML [3] and OpenMath [4] have been defined for embedding mathematical expressions in web documents. Up to now, all formulae had to be included as images with all the shortcomings of this approach. With these new developments, it is very likely that we will encounter mathematical web pages with MathML or OpenMath encoded expressions through the Math-Net Pages soon, e.g. as preprint. Thus, one goal in the near future is to be able to read and interpret these pages automatically. MathML and OpenMath both allow to encode the mathematical content of the expression.

MathML also allows to specify the presentation only, which might be particularly complicated, since the semantic content of the expression could eventually be lost in the course of using presentation markup. Unfortunately, presentation markup is most widely used and the problem to recover semantics from the presentation is a nontrivial one. One has to watch carefully which type of extra information regarding the semantics will be provided by the creators of MathML expressions such as Web editors or computer algebra systems.

4 Resume and Outlook

The experiences with Math-Net has shown that a community-driven and a community-based information and communications system is realistic. Such an information and communication system has to be distributed and open for participation within the whole mathematical community. To build up a successful, efficient, and long-term scientific information and communication system it is necessary

- to integrate the mathematical community

- to build up a personal and technical infrastructure for this system
- to develop the technical concepts and methods permanently basing on the Web developments and oriented at the needs of the users.

Of course, Math-Net as of today does not cover all relevant tasks and the information provided is still incomplete. But in general it seems to be doable. There are a lot of pretentious tasks for the future of digital information and communication in mathematics.

References

1. Committee on Electronic Information and Communication of the International Mathematical Union: Recommendations on information and communication, August 2002 (at ICM 2002), <http://www.ceic.math.ca/recommendations>, also available at <http://www.math-net.org/charter>
2. The Unicode Consortium: The Unicode Standard Version 3.0, Addison Wesley Longman Inc, 2000
3. World Wide Web consortium: MathML recommendation 2.0 at <http://www.w3.org/Math>
4. The Openmath standard (draft) The ESPRIT OpenMath Project, NAG Ltd, Oxford, UK , also at <http://www.openmath.org>