

The MSRI Digitization Project

What we want to do

The World Digital Mathematics Library has set as its goal the digitization of the entire past mathematical literature -- all of it, in electronic form and, eventually, available on the Internet. This project is a part of this larger effort meant to carry out this work on a large number of mathematics journals in a manner that makes the material maximally useful to all scientists and scholars. The project will also show how the WDML effort can work effectively *within* the established system of scholarly journals to increase their value, benefiting not only scientists but librarians and publishers as well.

Using funds from the Moore Foundation, the Mathematical Sciences Research Institute will award and coordinate grants for the purpose of digitizing complete runs of mathematics journals. The owners of the journals (referred to as "publishers", although they may not be traditional publishers) will receive the rights to incorporate the digitized material into their present holdings. In return, they must agree to make a substantial portion of the older material freely accessible on the Internet.¹ If an owner does not wish to post the digitized material, then the owner must agree to allow another organization to post the material under the same conditions. All material will be archived, and the entire effort will be carried out in cooperation with major research libraries.

The digitized material will consist of the following components:

- Bibliographic data (metadata) associated to each component of the journal
- High quality scanned images of each page; corresponding text files derived from optical character recognition; and software that allows searching (using these files).
- When present, a separate list of references with links to the corresponding items in *Math Reviews* and *Zentralblatt Math*.

This last part (the list of references with links) is a feature that is missing in previous digitization projects. It is the key to making this material more than merely a collection of files, providing a mechanism to connect all of mathematics on the Web.²

Why we want to do it

For scientists, the value of creating this web of scholarship is evident: A vast collection of the past mathematical literature will become available, at least in principle, to scholars in every part of the world. And because mathematicians depend crucially on past scholarship for their present work, this collection of material will be an indispensable tool, now and in the future.

¹ For some publishers, a substantial portion consists of all material more than 5 years old. The term "substantial portion" may mean something different for each publisher. On the other hand, it is the older literature that should be made freely accessible, and it should be large enough to be valuable to the mathematics community.

² The reviewing journals have made a commitment to add and to maintain links to original articles, providing a mechanism for linking from one article to the next that can be updated over time.

But this project is valuable for publishers of mathematics journals as well. The system of journals is under stress at the moment from many causes -- rising costs that outstrip rising library budgets, demands to invest in new technology for electronic products, and increased competition from journal substitutes, which even in the experimental stage weaken the system. Publishers are justifiably worried about the future of their journals.

Addressing these worries requires more than hand-wringing. Publishers must add value to their journals in order to protect them, and this means doing more than merely adding new "features" (most of which are seldom used). Each journal is part of a larger system -- a system of interconnected scholarship on which future generations of scholars rely. The most effective way to add value to an individual journal is to add value to the entire system by making the system even more connected and even more reliable.

This is a key purpose of the MSRI project. By digitizing large collections of journals and connecting them to each other (and the present mathematical literature), we connect more than a century of mathematical literature in a way no one could have imagined a century ago. We create a stable web of mathematics -- a web that mathematicians can navigate far more easily than ever before, and that mathematicians will quickly come to view as a critical resource for all their work.

If all this new material is only available to a few mathematicians, however, it adds little value to the system. The past literature must be accessible not just in principle but in fact. When one publisher makes older articles freely accessible, all other publishers (of current journals connected to those articles) benefit. If every publisher makes a substantial portion of the past literature freely accessible, every current journal becomes more valuable. The goal is to add as much value as possible to current journals, without eliminating incentives to support them through subscriptions. Publishers must work together to balance value and incentives in this way.

Creating a stable, reliable, and connected body of mathematical scholarship benefits every part of the research community -- publishers, libraries, and the scholars themselves.

When we want to do it

As soon as possible. There is urgency in moving forward because we need to protect the system of mathematical literature that has grown over the past centuries. Creating a stable environment before that system weakens is critically important. We have an opportunity now to begin that task.

March 21, 2005