Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takusstraße 7
D-14195 Berlin-Dahlem
Germany

M. GRÖTSCHEL, W. DALITZ, V. WINCZEWSKI

Math-Net: Ein Informationssystem der Mathematik

Preprint SC 98-25 (October 1998)

Math-Net: Ein Informationssystem der Mathematik

Martin Grötschel, Wolfgang Dalitz, Vincent Winczewski

30. Oktober 1998

Zwei Trends sind klar: Die Menge wissenschaftlicher Information wächst beständig, und Informationen werden zunehmend elektronisch bereitgestellt. Die Frage ist: Wie können wir damit „vernünftig“ umgehen?

In diesem Artikel beschreiben wir kurz, welchen Weg die Mathematik zur Bewältigung der elektronischen Informationsflut geht und wie sie versucht, hierbei Qualität zu sichern.

Mathematiker arbeiten traditionell mit Zeitschriftenartikeln und Büchern, heute vielfach mit Preprints (wegen der langen Verzögerung im Publikationsprozeß). Die erforderliche Information wird jedoch immer komplexer. Umfangreiche Software-Archive und Datensammlungen zu konkreten Anwendungsfällen helfen bei der Entwicklung von Lösungsalgorithmen, Datenbanken über mathematische Strukturen bei der Gewinnung von neuen Vermutungen. In Zukunft werden animierte Algorithmen, Videosequenzen und Multimediadokumente verstärkt die Lehre bereichern. Während gedruckte Zeitschriftenartikel und Bücher z.B. über die Datenbank MATH des Zentralblatts für Mathematik schnell und einfach aufzufinden sind, gibt es kein vergleichbares Suchsystem für elektronische Dokumente. Sammlungen elektronischer Zeitschriften (z.B. EMIS der European Mathematical Society unter http://www.emis.de/journals/) und Sammlungen von Links auf elektronische Zeitschriften (z.B. die Osnabrücker elib unter http://elib.uni-osnabrueck.de/MATHEMATIK/) sind im Aufbau, gleiches gilt für elektronische Preprints (LANL, MPRESS), bei den übrigen Dokumenten weiß man jedoch kaum, wie man sie finden soll.

Ein erster Ansatz ist meistens die Benutzung allgemeiner Suchmaschinen wie AltaVista oder HotBot. Sie liefern in gewissen Fällen gute Resultate. Man kann aber nicht gezielt fachspezifisch suchen und wird u.U. von riesigen Trefferlisten erschlagen. Außerdem: Wie steht es um die fachliche Qualität und Aktualität der gefundenen Dokumente?

Die Mathematik hat damit begonnen, ihre elektronischen Ressourcen auf fachspezifisch angemessene Weise zu erschließen. Dazu gehört die adäquate Bereitstellung elektronischer Dokumente verbunden mit dezentraler Qualitätssicherung. Hinzu kommen Vereinbarungen über technische, formale und organisatorische Standards sowie Verabredungen über die Zusammenführung dezentral erschlossener Daten in zentrale Komponenten. Die Umsetzung dieser Ideen erfolgt im Projekt „Informationsdienste für die Mathematik im Internet (Math-Net)“, das von DeTeBerkom über den DFN-Verein in der Zeit von Anfang 1996 bis März 1999 gefördert wird. Neun mathematische Fachbereiche nehmen unter der Leitung des Konrad-Zuse-Zentrums für Informationstechnik (ZIB) daran teil; insgesamt beteiligen sich derzeit rund 50 der etwa 70 mathematischen Einrichtungen in Deutschland.

Dezentrale Qualitätssicherung und personelle Infrastruktur

Rechner und Software können bestenfalls helfen, die formale Qualität von Dokumenten zu sichern. Uns kommt es aber mehr auf die inhaltliche Qualität an. Ein wichtiger Grundgedanke des Math-Net ist daher der Aufbau und langfristige Erhalt einer *personellen Infrastruktur*. Die Fachbereiche stehen für das auf ihren Servern bereitgestellte Angebot gerade. Sie haben ein originäres Eigeninteresse, daß Ihre Server einen hilfreichen, gut strukturierten und möglichst alle wichtigen Bereiche abdeckenden Service bieten. So wie es überall Bibliotheksbeauftragte gibt, gibt es an den Math-Net-Fachbereichen *Informationsbeauftragte*, die sowohl für die Kontinuität wie auch für die Qualität des Angebots verantwortlich zeichnen. Darüberhinaus sind auch übergeordnete Organisationen wie die Deutsche Mathematiker-Vereinigung (DMV), die Konferenz der mathematischen Fachbereiche (KMathF) und das Zentralblatt für Mathematik in die Aktivitäten eingebunden und betreiben in Kooperation mit den Informationsbeauftragten die Weiterentwicklung des Math-Net. Erste Kontakte bestehen zu weiteren europäischen, amerikanischen und internationalen Organisationen.

Strukturierung tut Not

Untersucht man heute die Einstiegsseiten universitärer WWW-Server, stellt man eine große strukturelle Inhomogenität fest. Wichtige Einstiegspunkte fehlen oft gänzlich (obwohl über indirekte Verweise zugänglich) oder tragen individuelle, nicht notwendig intuitiv verständliche Bezeichnungen. Wer hätte schon erwartet, daß das lokale Bibliothekssystem unter der Kategorie „Produkte der Einrichtung“ zu finden ist oder der Verweis auf die Preprints mal unter der Rubrik „Bibliothek“, mal unter „Mathematik Allgemein“ und dann wiederum unter „Forschung“ eingeordnet ist?

Ein naheliegendes Hilfsmittel zur ersten Orientierung auf WWW-Servern ist eine abgestimmte Strukturierung des Informationsangebotes. Dieser Gedanke wurde innerhalb des Math-Net-Projektes schon sehr früh aufgegriffen. Seine Realisierung bedurfte jedoch eines langen Weges der gegenseitigen Verständigung. Die Beteiligten einigten sich auf die Einführung von wenigen, überschaubaren Kategorien und Unterkategorien, die die relevanten Informationen gliedern und zusammenfassen. Das Ergebnis ist eine genormte WWW-Seite, genannt *Math-Net-Seite*, die alle beteiligten Institutionen als sogenannte *Secondary Homepage* (unter genormter URL) anbieten. Bild 1 zeigt die (englischsprachige Variante der) Math-Net-Seite der Technischen Universität München.

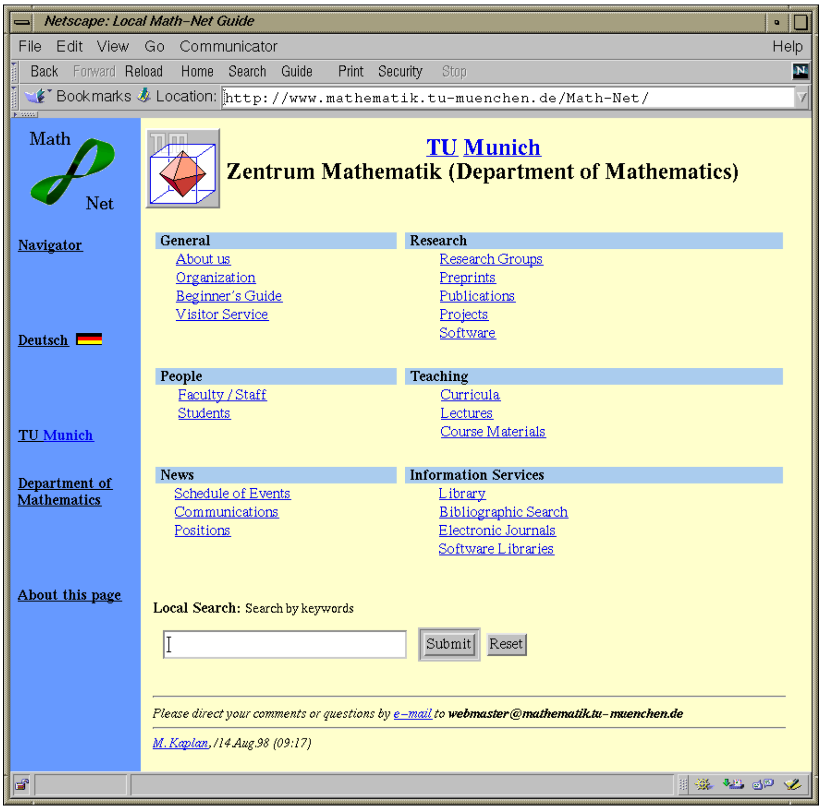

Bild 1

Aus ihr wird ersichtlich, daß man die Vielfalt des Angebots auf sechs Kategorien reduziert hat. Diese tragen intuitiv einleuchtende Bezeichnungen. Jede Kategorie ist in weitere Unterpunkte gegliedert. Auch deren Inhalt ist aus den Titeln erschließbar. Der Zugang zu einer lokalen Suchmaschine, die eine Volltextsuche auf dem lokalen Server ermöglicht, ergänzt die Math-Net-Seite. Es wurde ganz bewußt darauf verzichtet, alle Institutionen auf eine einheitliche Gestaltung ihrer „Primary Homepage“ zu verpflichten. Hier können sie weiterhin Individualität pflegen.

Metadaten und Suchmaschinen

Die Einführung der Math-Net-Seite ist als kurzfristig Maßnahme gedacht, um zielsicherer an relevante Informationen zu gelangen. Es bedarf sicherlich weiterer Maßnahmen, um in einem umfangreichen Angebot noch präziser suchen zu können.

Naheliegend ist die Verwendung von Metadaten, also die Bereitstellung von Daten über schon vorhandene Daten, wie es bei Bibliotheken schon lange Tradition ist. Frühzeitig ist von Seiten der Math-Net-Projektteilnehmer die Dublin-Core-Initiative begrüßt und favorisiert worden, die speziell die Entwicklung von Metadaten für elektronische Dokumente vorangetrieben hat. Das *Dublin-Core (DC) Metadatenset* zeichnet sich durch die Konzentration auf lediglich 15 wichtige Elemente aus, sieht aber Skalierbarkeit vor und bietet die Möglichkeit einer fachspezifischen Ausprägung. Besonders attraktiv erscheint die Verwendung der DC-Metadaten im Zusammenhang mit Suchmaschinen. Das *Harvestsystem* berücksichtigt z.B. solche Metadaten und bietet darüber hinaus entscheidene architektonische Vorteile, die der verteilt ausgerichteten Struktur des Math-Net entgegenkommen. Die in einer beteiligten Institution erfaßten Daten und Metadaten können zur lokalen Suche benutzt werden. Harvest ermöglicht zusätzlich die Weitergabe dieser Daten an andere Harvestsysteme, z.B. an übergeordnete Verbünde zur fachspezifischen Preprintsuche. Damit sind regionale oder globale Suchmaschinen (auch zu speziellen Themenkreisen) möglich. Bild 2 zeigt die Schnittstelle zur deutschlandweiten Suche nach mathematischen Preprints.

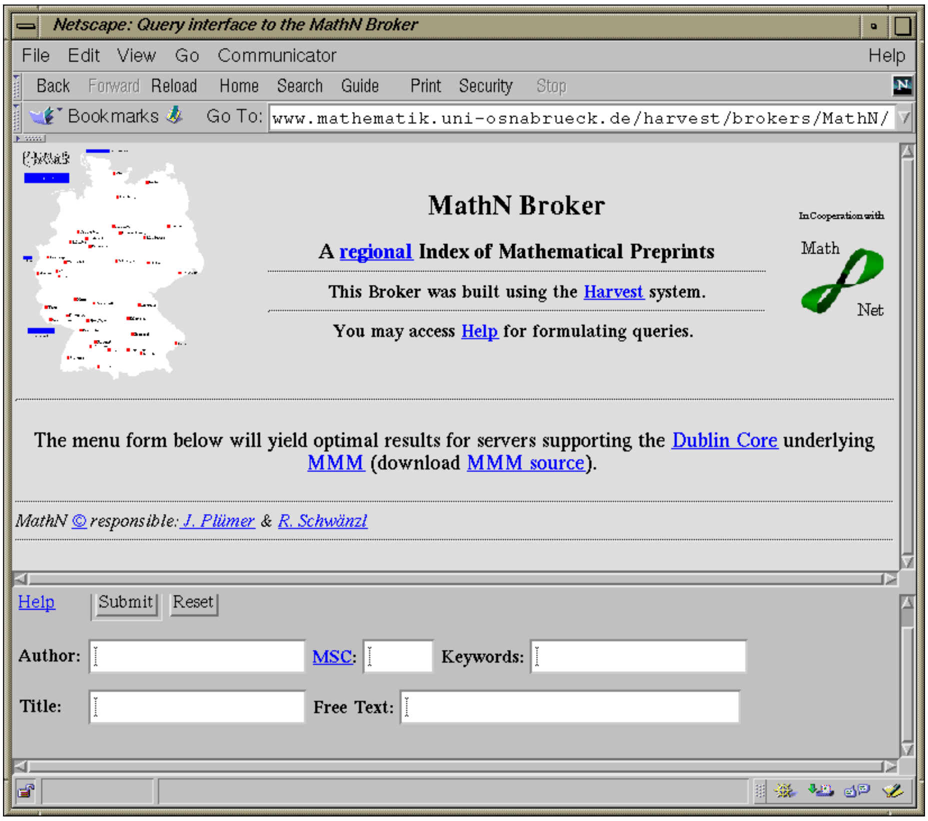

Bild 2

Dieser Service wurde inzwischen erweitert und verbindet unter dem Namen *MPRESS* weitere Preprint-Server aus Österreich, Frankreich und Kanada. Ein weltweiter Ausbau ist geplant.

Das Math-Net bietet auch Suchmöglichkeiten in einem Softwarearchiv. Ein auf Metadaten basierendes Personensuchsystem ist im Aufbau. Durch den dezentralen Ansatz des Systems ist damit auch dem Datenschutz Rechnung getragen, denn jede Institution ist für das rechtlich korrekte Angebot der eigenen Daten verantwortlich.

Die Flexibilität des Harvestsystems zeigt sich auch darin, daß nicht jede am Math-Net teilnehmende Institution ein eigenes System installieren muß. Mit einer sogenannte „Remote-Indexierung" lassen sich deren Daten einbinden.

Zentrale Komponenten dezentral pflegen: Zusammenführen, was zusammengehört

Nicht jede beteiligte Institution bietet (derzeit) alle Informationen und Dokumente elektronisch an. Hilfreich ist daher ein Überblick über die Einrichtungen, die z.B. Preprints offerieren. Ein Wunsch dieser Art führt häufig zur Installation eines neuen zentralen Dienstes. Die Konzeption und die Werkzeuge des Math-Net ermöglichen die Realisierung und Pflege eines solchen Dienstes mit relativ geringem Aufwand. Die Information, ob (und wenn dann wo) eine Institution „Preprints" oder „Hinweise für Studienanfänger" bereitstellt, kann aus der (genormten) Math-Net-Seite der Institutionen automatisch herausgefiltert und in einer Übersicht zusammengestellt werden.

Das Math-Net offeriert daher einen Dienst, genannt *Math-Net Navigator*, der aus einer geschickten Zusammenführung und übersichtlichen Darstellung der Unterpunkte der Math-Net-Seite besteht. „Klickt" man im Math-Net Navigator z.B. auf „Schedule of Events", so erhält man alle Institutionen des Math-Net, die einen mathematischen Veranstaltungskalender führen. Der Math-Net Navigator bietet außerdem Zugang zu lokalen und übergeordneten (auch regionalen) Suchmaschinen. Bild 3 zeigt den Beginn einer Liste von Institutionen des Math-Net, die Preprints anbieten.

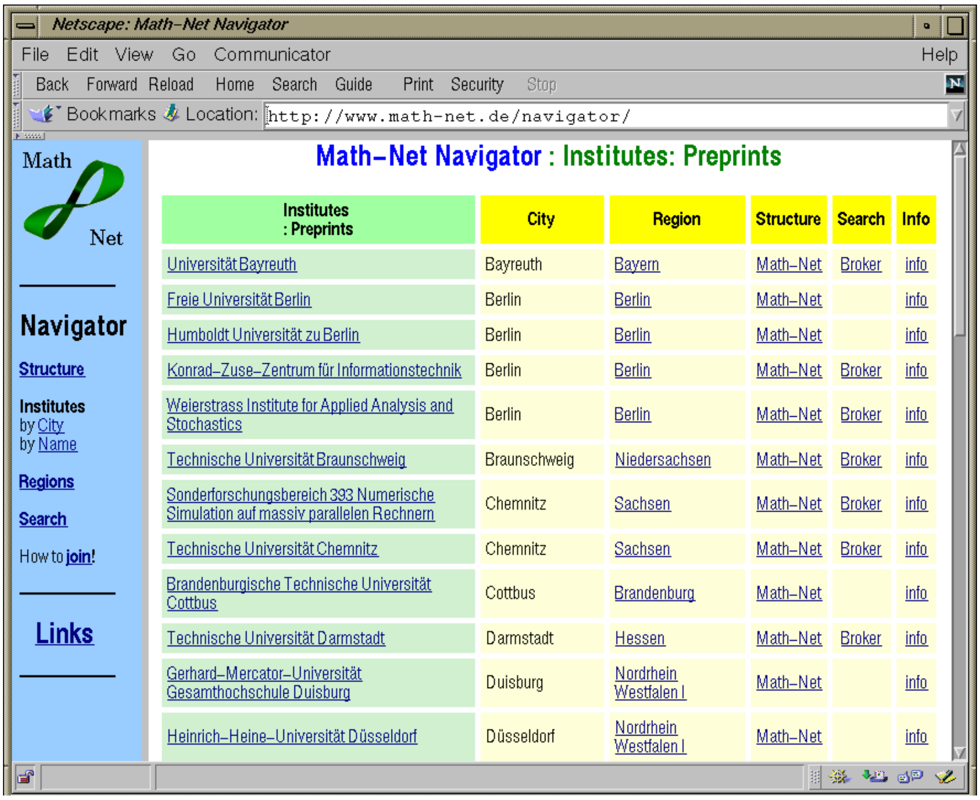

Bild 3

Durch Klicken auf einen hier angebotenen Link kommt man direkt zum Preprint-Verzeichnis der betreffenden Einrichtung.

Langfristig soll dieser Service durch Suchmöglichkeiten ergänzt werden, die mit Hilfe von Metadaten eine feinere Granulierung und Steuerung erlauben.

Nicht alles geht automatisch

Nicht alle weltweit zugänglichen mathematischen Ressourcen sind (mit vertretbarem Aufwand) automatisch erfaßbar. Daher ist in einer gesonderten Initiative (begonnen von Joachim Lügger, ZIB) eine Sammlung mathematikrelevanter Links erstellt worden: die *Math-Net Links to the Mathematical World* (Bild 4). Mit über 2000 aktuellen Verweisen stellt sie heute eine der größten Sammlungen in diesem Bereich dar. Das Spektrum reicht von elektronischen Journalen, über Softwaresammlungen, bis hin zu digitalisierten historischen Dokumenten und Materialien zur Geschichte der Mathematik.

Es versteht sich von selbst, daß diese Sammlung auch mit dem Harvestsystem erfaßt und damit absuchbar ist. Da es immer schwerer wird, solche Sammlungen aktuell zu halten, ist daraus ein verteilt organisiertes Autorensystem entstanden. Verschiedene Moderatoren haben bestimmte Teilbereiche übernommen und sind für deren Aktualität verantwortlich. Leser dieser Seiten tragen durch eigene Beiträge zur Vollständigkeit und zur Aktualität bei.

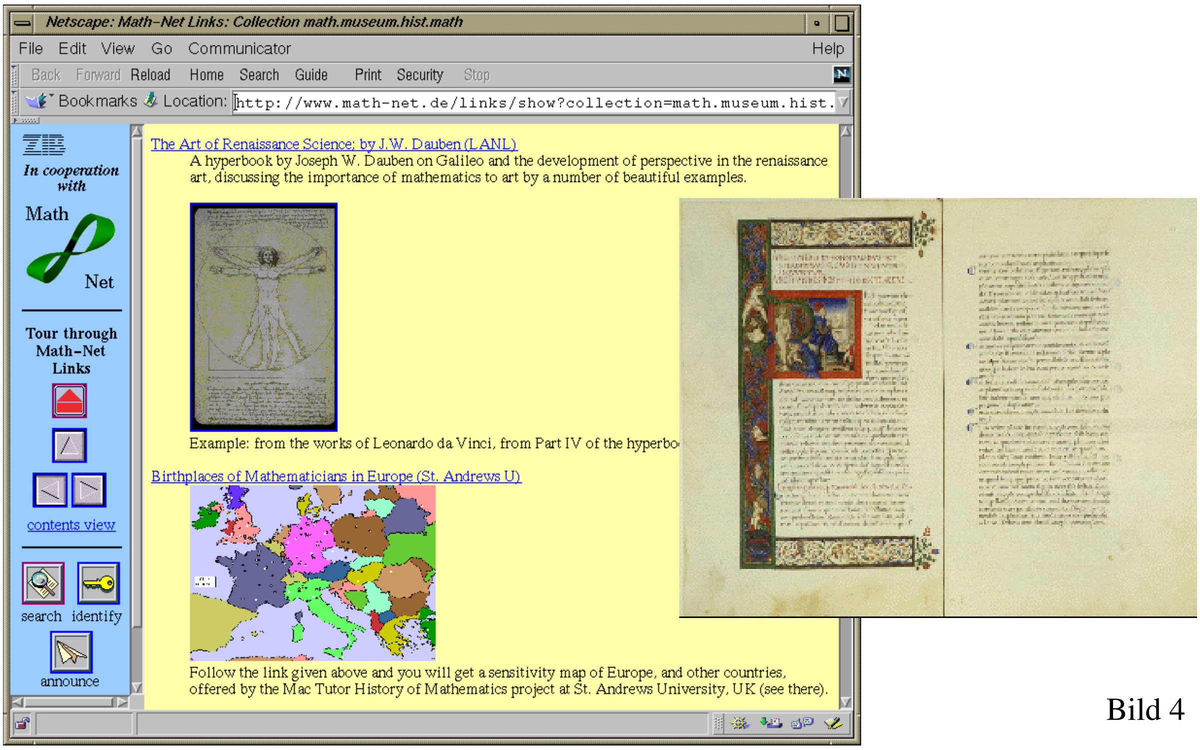

Bild 4

Ausblick

Die Mathematik verfolgt mit dem Math-Net das Ziel, eine technische und personelle Infrastruktur zu schaffen, die die Möglichkeiten moderner Informationssysteme ausnutzt. *Dezentral* und *kooperativ*, *offen* und an *internationalen Standards* ausgerichtet, sollen im Math-Net die wichtigsten elektronischen Ressourcen gebündelt und zugänglich gemacht werden.

Die *Internationale Mathematical Union* (IMU) hat im August 1998 beschlossen, ein *Committee on Electronic Information and Communication (CEIC)* zu gründen. Das Math-Net-Projekt wird seine Erfahrungen in das CEIC einbringen und seine Aktivitäten mit den internationalen Partnern abstimmen. Dadurch eröffnet sich eine realistische Chance, die Ideen des Math-Net Projekts in einem internationalen Verbund zu verwirklichen.

Kontakt

Weitere Informationen zum Math-Net erhalten Sie unter dem URL: ***http://www.math-net.de/***
oder bei Wolfgang Dalitz • Konrad-Zuse-Zentrum für Informationstechnik (ZIB) • Takustr. 7 • D-14195 Berlin • E-Mail: dalitz@zib.de • Tel.: (030) 841 85-201 • Fax: (030) 841 85-269