

Die WDML – Ideen und Wirklichkeit

Die World Digital Mathematics Library WDML verfolgt das Ziel, alle mathematischen Publikationen elektronisch zur Verfügung zu stellen. Mit der Electronic Library of Mathematics in EMIS (European Mathematical Information Service) trägt FIZ Karlsruhe zur Realisierung dieser Idee bei.

Bernd Wegner

1. Die Idee der WDML

Das mathematische Wissen wird durch die Gesamtheit der mathematischen Publikationen repräsentiert. Die mathematische Bibliothek ist das Laboratorium des Mathematikers. In der Erforschung neuer Themen und Problemstellungen sind die dazu bereits publizierten Ergebnisse zu berücksichtigen. Je besser der Zugriff auf die vorhandenen mathematischen Publikationen ist, desto komfortabler stellt sich das Umfeld für die mathematische Forschung und die Anwendung mathematischer Ergebnisse in Wissenschaft und Technik dar.

Angesichts der Vielzahl mathematischer Forschungsartikel ist es unmöglich, diese an jeder mathematischen Bibliothek gedruckt vorrätig zu halten. Selbst Bibliotheken, deren Auftrag es ist, alle mathematischen Publikationen systematisch zu sammeln, haben

Schwierigkeiten, diesem Auftrag vollständig nachzukommen. Hinzu kommt, dass das Interesse an mathematischen Publikationen praktisch keinem Verfallsdatum unterliegt. Statistiken, die die Zitate in aktuellen Publikationen auswerten, zeigen ein unerwartet langsames Nachlassen älterer Zitate.

Die prinzipielle Möglichkeit, digitale Versionen von Publikationen ins Netz zu stellen, eröffnete die Chance, eine virtuelle Variante einer umfassenden Mathematik-Bibliothek aufzubauen. John Ewing publizierte in 2001 eine Studie (Twenty Centuries of Mathematics: Digitizing and disseminating the past mathematical literature. www.ams.org/ewing/Twenty_centuries.pdf), die erste Probleme bei der Entwicklung einer solchen Bibliothek (nachfolgend WDML genannt) aufzeigte und als erste Abschätzung von einem Betrag von 100 Millionen US Dollar sprach, der erforderlich wäre, um sämtliche gedruckten Bestände zu retrodigitalisieren.

Mit diesem Report als Grundlage erhielt die Bibliothek der Cornell Universität für die Jahre 2002/3 Projektmittel von der NSF, um in einem relativ großen Rahmen eine Machbarkeitsstudie für die WDML durchzuführen. Das Ergebnis der Studie ist unter der Adresse www.library.cornell.edu/dmlib/ veröffentlicht. Die Weiterführung der Idee im Rahmen einer globalen Initiative wurde von der IMU (International Mathematical Union) übernommen, ohne jedoch größeren Einfluss auf das konkrete Geschehen zu haben.

Der Aufbau einer ersten Version der WDML erfolgte in Rahmen von Einzelprojekten, denen es gelang, Mittel für die Digitalisierung von Teilen der mathematischen Literatur einzuwerben. Immerhin hatte die Machbarkeitsstudie einige technische Standards entwickelt, die mehr oder weniger von den Projekten umgesetzt wurden. Aber fasst man die bisher retrodigitalisierten Publikationen mit den digital produzierten zu dieser



Abb. 12: Digitalisierte Arbeit in ERAM. Für den schnellen Datentransfer niedrig aufgelöst.

ersten Version zusammen, so erlebt der hoffnungsfrohe Nutzer eine erste Enttäuschung. Der einzige integrierte Zugang zu den verteilt vorliegenden digitalen Archiven wird durch Links über die Referenzdatenbanken Zentralblatt MATH und MathSciNet gewährleistet. Die Zugänglichkeit unterliegt verschiedenen Subskriptionsmodellen, so dass dem mittellosen Nutzer letzten Endes nur ein Teil der Bibliothek zur Verfügung steht.

2. Was soll zur WDML gehören?

Ein großer Teil der Machbarkeitsstudie befasste sich mit dem beliebten Problem, wer welche Dinge bestimmen durfte und welche Bestimmungen zu beachten sind. Das Hauptproblem, das vorher einer Lösung bedurfte, betrifft jedoch den Inhalt der WDML. Ohne eine Idee zur Strategie, welche Inhalte zu welchem Zeitpunkt in die WDML integriert werden sollten, gibt es keine geeignete Grundlage für administrative Maßnahmen. Angesichts der unterschiedlichen Vorstellungen der verschiedenen Eigner von Inhalten für die WDML ist die Diskussion einer zentralen Administration für die WDML im Moment ohnehin nur akademischer Natur.

Eine Wunschliste, was man gern in der WDML hätte, und damit verbunden, ein Maßstab, wie weit die WDML schon gediehen ist, hängt von verschiedenen Parametern ab.

Zeit: Wie weit gehen wir in die Vergangenheit zurück? Hier haben die Chinesen mit der Digitalisierung der alten chinesischen Mathematik eine Pionierrolle übernommen. Welche Publikationen werden als mathematisch bezeichnet? Diese Grenzziehung war schon immer ein Problem für die Referateorgane in der Mathematik. Hier gibt es einen fließenden Übergang zu den Anwendungsgebieten der Mathematik und die Lösung liegt in der Kooperation mit ähnlichen Projekten für andere wissenschaftliche Disziplinen.

Qualität: Soll eine auf die Forschung bezogene qualitative Grenze gesetzt werden? Nur Zeitschriften mit einem Begutachtungssystem oder auch Preprints, Reports, Mitteilungen und andere graue Literatur? Vielleicht auch Lehrbuchliteratur und Manuskripte? Soll eine Beschränkung auf mathematisch hochentwickelte Länder erfolgen? Nach den bisherigen Erfahrungen beantworten das die Geldgeber für die Retrodigitalisierung mittels der Entscheidung, was finanziert wird, und es sind auch keine Alternativen in Sicht.

3. Digitalisierungsprojekte

Die Digitalisierung von nur in gedruckter Form vorliegenden mathematischen Publikationen liefert den Kern der WDML. Elektronisch generierte Publikationen, die für eine Installation in einer virtuellen Mathematik-Bibliothek ohne eine nachträglich digitalisierte Version auskommen, gibt es erst seit der Mitte der 90er Jahre und zu damaliger Zeit nur in einem sehr geringen Umfang.

Größere Bestände wissenschaftlicher Zeitschriften wurden zuerst von JSTOR retrodigitalisiert. Die Auswahl erfolgte aus den Publikationen im angloamerikanischen Bereich und enthielt die wichtigsten mathematischen Zeitschriften. Größere Mathematik-Anteile wurden von den Projekten von Elsevier und Springer erfasst, in denen sie komplett ihre verlagseigenen Zeitschriften bearbeiteten. Zur Zeit werden diese Aktivitäten auf die Buchproduktion erweitert. Kleinere Verleger und wissenschaftliche Gesellschaften mit relevanten Angeboten von mathematischen Zeitschriften wie SIAM und LMS zogen nach. All diesen Anbietern ist gemeinsam, dass für den Zugang Gebühren erhoben werden.

Freien Zugang gewähren digitale mathematische Archive, deren Einrichtung mit staatlichen Mitteln oder Geldern von Stiftungen gefördert wurden. Hier müssen aus deutscher Sicht in erster Linie die durch die DFG unterstützten Aktivitäten des GDZ an der SUB Göttingen genannt werden, die mathematische Publikationen im Rahmen der Projekte ERAM und RusDML digitalisiert haben. Im Rahmen von ERAM wurde auch die Digitalisierung des Jahrbuchs über die Fortschritte der Mathematik vorgenommen, wobei es hier um eine

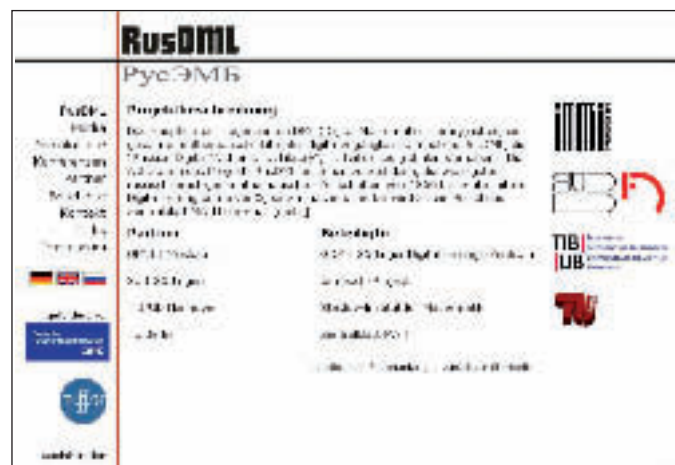


Abb. 13: Projektseite von RusDML

komplette Neuerfassung der Daten in Form einer Datenbank ging. Diese Daten wurden später in das Zentralblatt MATH integriert und werden dort einer fortlaufenden Nachbearbeitung unterzogen. RusDML hingegen behandelte russische Zeitschriften in Kooperation mit GPNTB Moskau als russischem Partner.

Die französischen Zeitschriften waren und sind Gegenstand der von Frankreich geförderten Arbeiten von NUMDAM in Grenoble. An diese Projekte haben sich eine Reihe von nationalen Digitalisierungsprojekten in Europa angeschlossen, so dass wir inzwischen nationale DMLs in Polen, Tschechien, Spanien, Serbien, Bulgarien, Portugal und der Schweiz haben. Weltweit kommen bemerkenswerte Aktivitäten in China, Korea und den USA hinzu. KISTI hat schon seit einiger Zeit sämtliche mathematische Zeitschriften digitalisiert, die in Korea herausgegeben werden. In China wurden im Rahmen eines groß angelegten Projekts neben den Zeitschriften auch Bücher und die alte chinesische Mathematik komplett erfasst. Der Zugang zu diesem Archiv wurde von der Bibliothek an der Tsinghua Universität in Beijing eingerichtet. In den USA hat die Bibliothek an der Cornell Universität relevante Bestände an digitalisierten mathematischen Publikationen aufgebaut.

Es gibt ferner einige mathematische Zeitschriften, die die Digitalisierung ihrer alten Bestände aus ihrem Budget vorgenommen haben. Nimmt man all diese Aktivitäten zusammen, so hat sich ein Teil der WDML bereits in Form eines Flickenteppichs entwickelt. Durch die Verlinkung der Bestände mit ihren Einträgen im Zentralblatt MATH hat sich jedoch eine erste Form eines integrierten Zugriffs ergeben. Diese Koppelung sowie die allgemeinen Diskussionen zur WDML haben ferner dazu geführt, dass einige gemeinsame Standards erkennbar sind.

4. EMANI und ELibM

Gute Kooperationsansätze sind in der EMANI-Gruppe zu erkennen, die aus den Bibliotheken in Göttingen, an der Cornell Universität, an der Tsinghua Universität, NUMDAM in Grenoble und dem Springer-Verlag besteht. EMANI steht für Electronic Mathematics Archiving Network Initiative und wurde in den vergangenen Jahren vom Verfasser dieses Artikels koordiniert. Die an den Bibliotheken und bei NUMDAM digitalisierten Bestände wurden in das System eingebracht, und eine der momentan wahrgenommenen Projektaufgaben besteht darin, hierfür einen integrierten Zugriff einzurichten.



Abb. 14: Projektseite von EMANI

Der Zugriff soll jedoch alle digitalen Angebote von Mathematik bei den EMANI-Partnern erreichbar machen. Zusätzlich zur Produktion bei Springer betrifft das die elektronisch produzierten Zeitschriften im Projekt Euclid an der Cornell Universität sowie den Spiegel der ELibM an der SUB Göttingen. Ferner kümmert sich EMANI um ein wichtiges Folgeproblem bei dem Angebot elektronischer Publikationen, der Langzeit-Archivierung elektronischer Dokumente. Es geht hier vorwiegend um den langfristigen Erhalt der Lesbarkeit der Dokumente, also auch um ein generelles Problem für die WDML. Zur Zeit werden im Rahmen von EMANI die Werkzeuge hergestellt, die die Langzeitarchivierung in einem verteilten System ermöglichen. Für den Archivierungsprozess selbst soll auf inzwischen anderweitig eingerichtete Verfahren zurückgegriffen werden.

Der zweite Teil der WDML besteht aus den elektronisch generierten mathematischen Publikationen, die durchaus noch in gedruckter Form vorliegen können, aber auf jeden Fall eine Webversion anbieten. In den 90er Jahren sind die meisten mathematischen Zeitschriften auf diese Form des Angebots umgestiegen. Damit haben Zeitschriften, die schon länger erscheinen, sowohl frühere Anteile im retrodigitalisierten Bereich als auch spätere elektronisch generierte Bestände. Da die Autoren ihre Manuskripte inzwischen ausschließlich mit Hilfe von Computern schreiben oder schreiben lassen, ist die Herstellung einer elektronischen Fassung einer Arbeit im Fall der Annahme zur Publikation kein schwieriges Unterfangen mehr.

Um die Sichtbarkeit im Web zu verbessern (und natürlich auch Vorteile im Marketing zu erzielen), haben die Verlage die elektronischen Versionen ihrer Zeitschriften gebündelt und für diese Bündel spezielle Zugangs-

strukturen eingerichtet. Der Flickenteppich der WDML wird auf dieser Ebene leicht umgeordnet. Einzige globale Zugangsstruktur bleiben wiederum nur die beiden Referateorgane Zentralblatt MATH und MathSciNet.

Einzelne Zeitschriften, die von einem kleinen Verlag, einer Akademie oder einer mathematischen Gesellschaft publiziert werden, sitzen hier zwischen den Stühlen und verlieren an Popularität. Dass ihre Produktion unter den neuen Bedingungen wesentlich preiswerter erfolgen kann und sie damit auch weniger teuer sind, hilft ihnen nicht viel weiter. Diese Situation wurde schon in der ersten Hälfte der 90er Jahre von den Initiatoren von EMIS (European Mathematical Information Service, www.emis.de) auf europäischer Seite und von der Bibliothek der Cornell Universität für Nordamerika erkannt, als diese mit der Planung vom Projekt Euclid begann. Euclid hat inzwischen die Produktion einer Reihe von Zeitschriften übernommen, die vorher auf sich selbst angewiesen waren, und diese können zu moderaten Preisen abonniert werden.

In EMIS, einer Kooperation von FIZ Karlsruhe mit der Europäischen Mathematischen Gesellschaft EMG, wurde die Electronic Library of Mathematics ELibM eingerichtet, in der frei zugängliche mathematische Publikationen elektronisch generiert oder elektronische Auflagen von externen Anbietern gespiegelt werden.



Abb. 15: Der Eingang zur ELibM

EMIS selbst hat ein weltweites System von ca. 40 Spiegeln eingerichtet und leitet Kopien dieser Inhalte an alle Spiegel weiter. Mit dieser Bündelung und Verteilung ist für eine verbesserte Popularität und Zugänglichkeit der Inhalte der ELibM gesorgt, und nebenbei genügt das System einem wichtigen Sicherheitsaspekt, indem mehrere Kopien der ELibM unabhängig voneinander vorrätig gehalten werden.

Der Schwerpunkt der ELibM liegt im Bereich der Zeitschriften. Nachdem von den Organisatoren am Anfang viel Überzeugungsarbeit geleistet werden musste, um Zeitschriften von dem Vorteil der freien Zugänglichkeit im Web zu überzeugen, war bald eine kritische Masse aufgebaut worden, die von selbst überzeugt. Zur Zeit liegt der Stand der Zeitschriftenabteilung bei fast 100 Zeitschriften, die ca. 15 Gigabyte an Speicherplatz benötigen. Neue Interessenten melden sich jetzt von selbst. Es sollte jedoch erwähnt werden, dass bei Zeitschriften, die von Erlösen aus Abonnements abhängen, eine Schutzperiode („moving wall“) von in der Regel drei Jahren eingerichtet wurde. Der freie Zugang zu den Artikeln wird erst nach der Schutzperiode geschaltet. Die „Nähe“ der ELibM zum Zentralblatt MATH ermöglicht eine schnelle und zuverlässige Verlinkung zwischen beiden Diensten, was zusätzlich den Zugriff auf die ELibM erleichtert.

Angesichts der Heterogenität der verschiedenen Teile der bisher aufgebauten Bestände für die WDML, oder besser gesagt, der Bestände, die prinzipiell zu einer WDML gehören sollten, erscheint es sehr schwierig, den Idealzustand der WDML erreichen zu können. Hinzu kommt, dass es unterschiedliche Interessen und Erlöserwartungen bei den Eignern der Bestände gibt. Dieser Idealzustand wird durch einen integrierten und vor allem offenen Zugang charakterisiert, den sich auch Mathematiker in Entwicklungsländern leisten können. Weitere Diskussionen auf der Ebene der IMU könnten Druck ausüben, so dass sich in dieser Richtung etwas bewegt. Die Verlinkung mit Referenzdatenbanken wie dem von FIZ Karlsruhe herausgegebenen ZMATH wird auf absehbare Zeit den einzigen integrierten Zugang zu den Beständen der WDML liefern.

Prof. Dr. Bernd Wegner
 Chefredakteur Zentralblatt MATH