



Niedersächsische Staats- und
Universitätsbibliothek Göttingen



Archivierungsstrategien in der Mathematik

Dr. Thomas Fischer
Metadaten und Datenbanken
SUB Göttingen
fischer@mail.sub.uni-goettingen.de



Protokolle -
Portale - Profile



Archivierung in der Mathematik

- Warum?
- Was?
- Beispiele: Jahrbuch-Projekt, EMANI, DML
- Strategien
- Standards: Allgemein
- Dateiformate
- Metadaten
- Archive
- Verantwortung

Warum?

- ▶ Bücher und Mikroformen werden von Bibliotheken „archiviert“
- ▶ Digitale Bestände wachsen:
 - Aus Retrodigitalisierung
 - Originär digitale Dokumente
- ▶ Dabei entstehen Formen in großer Vielfalt:
 - Texte (Word, PDF, TeX, DVI, Postscript ...)
 - Bilder (TIFF, GIF, JPEG ...)
 - Film (MPEG, mov, wav ...)
 - Interaktive Formate (?)
 - Multimedia allgemein

Aber: Keine geklärte Archivpolitik

Verantwortung, Ort, Methode der Archivierung sind ungeklärt!

Was?

- ▶ Historische Daten: Retrodigitalisierung
 - Jahrbuchprojekt (Deutschland)
 - Distributed Digital Library of Mathematical Monographs (USA-Deutschland)
 - DML: Digital Mathematics Library (USA-Europa)
- ▶ Aktuelle Daten: Digitale Mathematische Quellen
 - Elektronische Zeitschriften
 - Elektronische Dissertationen und Habilitationen
 - Monographien
 - Preprints
- ▶ Ziel: Den Wissensschatz der Mathematik vollständig elektronisch verfügbar zu machen und zu erhalten!

Beispiel Jahrbuchprojekt

Aufgaben:

- ▶ Digitalisierung des „Jahrbuch über die Fortschritte der Mathematik“ (1868 -1942) als Datenbank (JFM database) mit Abstimmung mit Zentralblatt
- ▶ Digitalisierung der wichtigsten referierten Mathematischen Werke (ca. 20%)
- ▶ Nachweise der Dokumente über die Datenbank und Präsentation im Internet über Webbrowser mit Standard-Technik

Ergebnis:

Bis jetzt über 1,2 Mio. Seiten

Beispiel DML

Aufgaben:

- ▶ Digitalisierung des gesamten Mathematikbestandes
- ▶ Kooperation USA (NSF) und Europe (EU-RP6)
- ▶ Koordination mit bestehenden und parallelen Initiativen, Abgrenzung der Aufgaben ...
- ▶ Geschätzt: 50 Mio. Seiten, Kosten 100 Mio. US-\$
- ▶ Standardisierung:
 - Metadaten (für Recherche, Verwaltung, Pflege etc.)
 - Datenformat (zukunftsicher)

Beispiel EMANI

EMANI =

Electronic Mathematics Archiving Network Initiative

▶ Partner:

- Springer-Verlag
- Universitätsbibliotheken:
 - Cornell (USA)
 - Tsinghua (Peking)
 - Göttingen

▶ Ziel:

- Verlässliches, allgemein zugängliches digitales Archiv
- Damit Garantie der dauerhaften Verfügbarkeit elektronische Dokumente

Strategien:

- ▶ Standards!
- ▶ Standards!
- ▶ Standards!
- ▶ ...

Standards, Standards...

- ▶ Archivierungs-Technik
 - Speichermedien und Zugriff
 - Langzeitfragen: Haltbarkeit, Migration, Transformation
- ▶ Standards für Datenformate
 - Textformate für Archivierung und Präsentation
 - Bildformate
 - Multi-Media
- ▶ Standards für Metadaten
 - Dublin Core (OCLC)
 - METS (Metadata Encoding and Transmission Standard, LoC)
 - Application Profile (z.B. DC-Lib AP)
- ▶ Standards für Archive bzw. Archivsysteme
 - Open Archival Information System (OAIS)
 - Trusted Digital Repositories
 - Langzeitverfügbarkeit digitaler Dokumente

Standards: Dateiformate

- ▶ Vielfältige Dateiformate erschweren Bewahrung und Bereitstellung
 - Formate benötigen spezielle Programme
 - Programme benötigen spezifische Betriebssysteme
 - Betriebssysteme benötigen spezielle Hardware
 - ...
- ▶ Mögliche Auswege:
 - Emulation
 - „Hardware-Museum“
- ▶ Besser:
 - Offene Dateiformate
 - Public-Domain-Software
- ▶ Sicherere Formate
 - Mark-Up für Text (XML, TeX,...)
 - Für Bilder TIFF

Standards: Metadaten

Für den Austausch, die Verwaltung und das Auffinden der Dokumente sind Metadaten unabdingbar.

- ▶ Internationaler Standard: Dublin Core
- ▶ Im Detail nicht ausreichend, erweitert um weitere Datenfelder aus anderen Definitionsbereichen
- ▶ Verbindung verschiedener Metadatenstandards: Application Profiles
- ▶ Dazu Arbeitsgruppe MMI-DC bei Europäischer Normungsbehörde (CEN) zur Entwicklung einer Handreichung zu Application Profiles.

Standards: Archive

- ▶ Standard: Open Archive Information System
(Consultative Committee for Space Data Systems, 2002)
 - Prinzipien für die Organisation eines Archivs
 - Rahmenstruktur für Aufnahme, Speicherung, Verwaltung, Zugang und Lieferung
 - keine Implementierung
- ▶ Verlässliche Archive:
Trusted Digital Repositories (RLG/OCLC)
 - Qualitätssicherung für Archive
 - Klärung von Verantwortlichkeiten
 - “a framework of attributes and responsibilities for trusted, reliable, sustainable digital repositories capable of handling the range of materials held by large and small research institutions.”

Archiving responsibility

Best Current Practices: Recommendations on Electronic Information Communication (IMU, AMS):

- ▶ Traditionally, maintaining the older literature has been the responsibility of librarians rather than publishers. Even in the electronic age, scholars and the librarians who represent them have the greatest motivation among all of the affected parties to ensure the preservation of older material.
- ▶ We recommend that electronic archives of the mathematical literature should ultimately be under the control of the academic community.
- ▶ Langzeitverfügbarkeit digitaler Dokumente
 - Workshop 29./30. Oktober 2002 an Der Deutschen Bibliothek



Fragen, Anmerkungen ... ?

Vielen Dank für Ihre Aufmerksamkeit!

Dr. Thomas Fischer
Metadaten und Datenbanken
SUB Göttingen
fischer@mail.sub.uni-goettingen.de

