

Vernetzte Informationsdienste für die Mathematik

Wolfram Sperber

FIZ Karlsruhe, Zentralblatt für Mathematik

DINI, Vernetzungstage 2011

3.3.2011

Universität Osnabrück

Agenda

- Vorbemerkung
- Informationsdienste in der Mathematik – ein Rückblick
- Die Herausforderungen
- Die Aktivitäten des Zentralblatt
- Ein Fazit

Vorbemerkung

- Interdisziplinarität: besondere Rolle und Stellung der Mathematik
- Schlüsseltechnologien heute: „Math Inside“
- breites Anwendungsspektrum der Mathematik, d.h. Mathematische Modelle und Methoden kommen in allen Wissenschaften, aber auch in Industrie, Dienstleistung und im Management zum Einsatz

Informationsdienste in der Mathematik: – das Jahrbuch über die Fortschritte der Mathematik (JFM)

- 1868: Gründung des JFM - erster eigenständiger Informationsdienst der Mathematik
- Mathematik und Anwendungsgebiete (Mechanik, Physik, Optik, Astronomie, ...)
- Ziele: Erschließung und systematische Darstellung des mathematischen Wissens, Werkzeug für die Wissenschaftler (Suche und Überblick)

Jahrbuch über die Fortschritte der Mathematik (II)

- Idee Konzept) und Initiative für JFM gingen von der mathematischen Community aus
- Inhaltliche Erschließung:
 - Zuordnung der Arbeit zu einem mathematischen Gebiet
(über Kapitelstruktur)
 - Review: Besprechung der Arbeit durch Experten
(inkl. Verweise)

Das „Zentralblatt für Mathematik und ihre Grenzgebiete“

- 1931 – weiteres Referatejournal für die Mathematik und ihre Anwendungsgebiete
- Interdisziplinärer Charakter mit dem Fokus Mathematik wird schon im Titel hervorgehoben
- heute Datenbank ZBMATH (umfasst JFM):
Stand: mehr als 3.000.000 Nachweise mathematik-relevanter Publikationen seit 1828 bis heute

Interdisziplinarität in ZBMATH

- breites Anwendungsspektrum: andere Wissenschaften, aber auch Produktion, Logistik, Management, Kunst, ...
Basis: \sim 3.400 wissenschaftliche Journale
+ Bücher aus verschiedenen wissenschaftlichen Disziplinen (natürlich mit dem Fokus Mathematik)
- falls Arbeiten auch Anwendungsaspekte enthalten, duale Erschließung:
mathematische + Anwendungsaspekte

Interdisziplinarität: Klassifizierung

- Mathematical Subject Classification 2010:
 - 16 [von 63] Top-Klassen der MSC sind „reine“ Anwendungsgebiete
 - darüber hinaus gibt es in vielen MSC Top Klassen zusätzlich Unterklassen für die Anwendungen
 - andere Klassifikationen (PACS, UDC, ...) werden mitgeführt

Klassifizierung in der Zukunft?

Interdisziplinarität: Keywords, Verlinkung

- feinere Erschließung interdisziplinärer Inhalte von Publikationen durch zusätzliche Keywords aus den Anwendungsbereichen
- Verlinkung: Extraktion der Literaturreferenzen in den Publikationen (so weit wie möglich Verknüpfung mit Volltexten)

Mathematisches Wissen heute

- mathematisches Wissen liegt heute nicht mehr nur in Form von Publikationen vor
- mit der Entwicklung des Computers ist mathematische Software als neue Facette und als Brückenglied zu den Anwendungen der Mathematik entstanden

Herausforderungen für wissenschaftliche Informationsdienste (I)

- Anforderungen und Chancen an die Informationsdienste:
 - erweiterte und qualitative bessere Erschließung
 - bessere Vernetzung
 - Enhanced Retrieval
- Erweiterung um neue Inhalte und Formen, in der Mathematik, z.B. Software

Herausforderungen für wissenschaftliche Informationsdienste (II)

- Vernetzung sowohl innerhalb einer Wissenschaft als auch interdisziplinär
- Vernetzung - natürliche Fortsetzung und Weiterentwicklung der bestehenden Informationsdienste
- Rahmenbedingungen: Und alles möglichst automatisch, keine zusätzlichen Ressourcen

Zentralblatt Aktivitäten: MSC

- Verbesserung des bisherigen Angebots
- Verbesserung der Werkzeuge:
 - MSC (Stand):
Überführung der MSC in das SKOS Format
 - „Semantisches“ MSC:
Relationen, Definition der Klassen, Typisierung der Klassen

Zentralblatt Aktivitäten: Kontrolliertes Vokabular

- Aufbau eines kontrollierten Vokabulars für die Beschreibung und Erschließung mathematischer Publikationen
(wichtig für die inhaltliche Präzisierung der MSC, Perspektive: Vernetzung mit Vokabularen anderer Disziplinen)
- Techniken des automatisches Lernens
Trainingsmenge: MSC Vokabular, Keywords der Datenbank ZBMATH,

Zentralblatt Aktivitäten: Keyword-Analyse

- automatische Keywordextraktion mit Hilfe des kontrollierten Vokabulars
- Vorteile:
 - Kontextbezug (Trennung von mathematischen und Anwendungsvokabular)
 - Standardisierung der Keywords
 - für automatische Klassifizierung einsetzbar
 - inhaltliche Erschließung unterhalb der MSC-Ebene

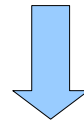
Zentralblatt Aktivitäten: Verlinkung

- Autorenidentifikation (Autorenprofile)
- Extraktion der Referenzen und Verknüpfung mit Volltexten
derzeit EuDML Projekt („European Digital Mathematics Library“)
- Verknüpfung mit grauer Literatur prinzipiell möglich

Zentralblatt Aktivitäten: Formelanalyse

- mit XML-Formaten sind Voraussetzungen da, um die inhaltliche Erschließung durch Formelanalyse zu verbessern

Identifikation der mathematische Terme und Formeln
+ Kontextanalyse



semantische Interpretation der Terme und Formeln
(Projektidee)


Zentralblatt Aktivitäten: SMATH-Projekt (I)

- Entwicklung und Aufbau eines Informationsdienstes für (mathematische) Software
- SMATH und Vernetzung:
mathematische Software hat Produktcharakter,
unmittelbares Werkzeug für Anwender
mathematische Software – Bindeglied zwischen
Mathematik und Anwendungen

Zentralblatt Aktivitäten: SMATH-Projekt (II)

- Schwierigkeiten: Beschreibungsmodell für Software
 - Inhaltliche Metadaten
(Subject [Granularität: Allgemeine Systeme versus Spezielle Systeme], Keywords, ...)
 - Technische Parameter (Betriebssysteme, Programmiersprachen, ...)
 - Versionierung
 - Lizenzbedingungen

Zentralblatt Aktivitäten: SMATH-Projekt (III)

- Schwierigkeiten (Rahmenbedingungen):
weitgehend automatisiert
- Besonderheit unseres Ansatzes: enger Bezug zwischen
mathematischer Literatur und Software
(Algorithmen werden in Publikationen dargestellt:
Wissenskette
Publikationen  Software)

Zentralblatt Aktivitäten: SMATH-Projekt (IV)

- ZBMATH-Daten als Ausgangsbasis für die Suche nach mathematischer Software:
Identifizierung von mathematischer Software +
Indikatoren für die inhaltliche Erschließung
(MSC-Klassifikation)
Einstieg: Suche nach bekannter Software +
Schlüsselwörter in ZBMATH

Fazit (I)

- Informationsprovider in den Wissenschaften stehen vor großen Herausforderungen
- Vernetzung ist zentrale Aufgabe und hat viele Facetten:
 - Interdisziplinarität,
 - Vernetzung unterschiedlicher Typen von Informationen (Metadaten mit Volltext einer Publikation, Publikation mit Software, Klassifikationssystem mit kontrolliertem Vokabular, ...)

Fazit (II)

- Vernetzung geht einher mit einer höheren Qualität der Informationsdienste und Mehrwert für den Nutzer (ermöglicht einen individuellen Zugang zu den Informationen)
- mit dem Semantic Web ist eine konzeptionelle und technische Basis für die Vernetzung von Informationen und Informationsdiensten gegeben

Fazit (III)

- aber: Vernetzung von Informationen und Informationsdiensten wird sich nicht im Selbstlauf vollziehen, diese müssen (inhaltlich und technisch) aufeinander abgestimmt und koordiniert werden, nicht zuletzt erfordert das auch finanzielle Mittel (Projekte)

Danke