



## **Minisymposium 29 - Information, Kommunikation und Bibliotheken für die Mathematik**

### **Kann man automatisch klassifizieren? Probleme und Ansätze automatischer Klassifikation**

WOLFRAM SPERBER (BERLIN)

Gegenwärtig erscheinen jährlich mehr als 70.000 mathematische Publikationen. Die beiden führenden Referateorgane in der Mathematik, Zentralblatt Math und Mathematical Reviews, klassifizieren diese Arbeiten gemäß der Mathematical Subject Classification bis heute manuell. Kann man mathematische Publikationen automatisch klassifizieren? Das immense Wachstum des Web und die Entwicklung leistungsfähiger Suchmaschinen haben die automatische Klassifikation bzw. Clustering zu einem vieldiskutierten Thema gemacht. Ein Ansatz zur automatischen Klassifizierung ist die Textkategorisierung. Das übliche Vorgehen der Textkategorisierung besteht darin, den Text in seine Bestandteile (Wörter oder Wortgruppen) zu zerlegen und anhand der Wörter bzw. Wortgruppen die Arbeit zu klassifizieren. Der Vortrag diskutiert verschiedene Ansätze der Textkategorisierung und deren Anwendung auf die Klassifizierung mathematischer Publikationen.