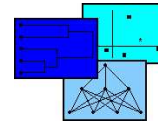


29. Jahrestagung der Gesellschaft für Klassifikation e.V.

---

**29<sup>th</sup> Annual Conference of the  
German Classification Society**



# **From Data and Information Analysis to Knowledge Engineering**

March 9–11, 2005

## **Program**



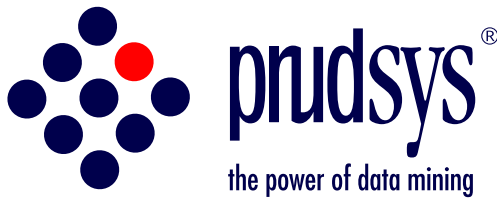
Otto-von-Guericke-University of Magdeburg

---





We would like to thank our sponsors:



4 Acknowledgments

- DaimlerChrysler AG
- Deutscher Sparkassen- und Giroverband
- Heins+Partner GmbH
- Otto-von-Guericke-Universität Magdeburg
- prudsys AG
- Springer-Verlag GmbH
- SPSS GmbH Software
- Stadt Magdeburg

# Table of Contents

<b>General Information</b> .....	<b>7</b>
Greetings .....	9
Organization .....	10
Conference Location .....	11
Information for Speakers .....	14
Registration Fees .....	14
Special Events .....	15
Social Program .....	16
Exploring Magdeburg .....	18
Public Transportation .....	20
Restaurants and Bistros .....	21
<b>Program</b> .....	<b>23</b>
Timetables .....	25
Doctoral Workshop .....	28
Detailed Conference Program .....	29
Wednesday, March 9 .....	29
Thursday, March 10 .....	38
Friday, March 11 .....	46
Librarians' Program .....	53



## General Information



## Dear conference participants

Welcome to the 29<sup>th</sup> Annual Conference of the German Classification Society, GfKI 2005 at the Otto-von-Guericke-University Magdeburg, under the theme “From Data and Information Analysis to Knowledge Engineering.” The GfKI 2005 is organized in cooperation with the Slovenian Artificial Intelligence Society (SLAIS). A rich and multi-faceted scientific program is expecting you. The next pages will provide you with an overview of the scientific events and with useful information on the conference and its social program.

We would like to draw your attention to several events that take place as part of the GfKI 2005: Starting this year, the best papers of the previous conference will be announced and awarded at the opening session. A Data Mining Competition was organized for the second time; the best solutions will be awarded during the reception and presented in the opening session (winner) and in a dedicated session on the second day of the conference (second winners). A Doctoral Consortium for PhD students has been organized for the first time on March 8, 2005, just before the conference. A poster session is scheduled for the second conference day, before the GfKI general assembly. An industrial track with two sessions on industrial applications and solutions is also taking place on March 10. The Librarians Workshop takes place in parallel to the conference on March 9 and 10.

We hope that you will enjoy the conference and wish you a pleasant stay in Magdeburg.

With best regards,

Myra Spiliopoulou

(Chair of Local Organization Team, University of Magdeburg)

Wolfgang Gaul

(Program Committee Chair, University of Karlsruhe)

Dunja Mladenić

(SLAIS, Josef Stefan Institute)

Werner Esswein

(Doctoral Consortium, TU Dresden)

Hans-J. Hermes

(Librarians Workshop, TU Chemnitz)

## Organization

### Local Organizers at the Otto-von-Guericke-University of Magdeburg

- Myra Spiliopoulou (Local Organization Chair)
- Rudolf Kruse (Publication Chair)
- Andreas Nürnberger (Publicity Chair)
- Christian Borgelt (Abstract Submission and Booklet Preparation)

### Program Committee

- Wolfgang Gaul (University of Karlsruhe)      (Chair)
- Hans-Hermann Bock (RWTH Aachen)
- Reinhold Decker (University of Bielefeld)
- Bernard Fichet (University of Aix-Marseille II, France)
- Rudolf Kruse (Otto-von-Guericke-University of Magdeburg)
- Hans-Joachim Lenz (Free University of Berlin)
- Dunja Mladeníć (J. Stefan Institute, Slovenia)
- Katharina Morik (University of Dortmund)
- Otto Opitz (University of Augsburg)
- Myra Spiliopoulou (Otto-von-Guericke-University of Magdeburg)
- Maurizio Vichi (University of Roma — “La Sapienza”, Italy)
- Claus Weihs (University of Dortmund)
- Klaus-Dieter Wernecke (Charité Berlin)

## Conference Location

Otto-von-Guericke-University of Magdeburg

Most of the main conference program takes place at:

Building 22A  
Pfälzer Straße  
39104 Magdeburg

Additionally:

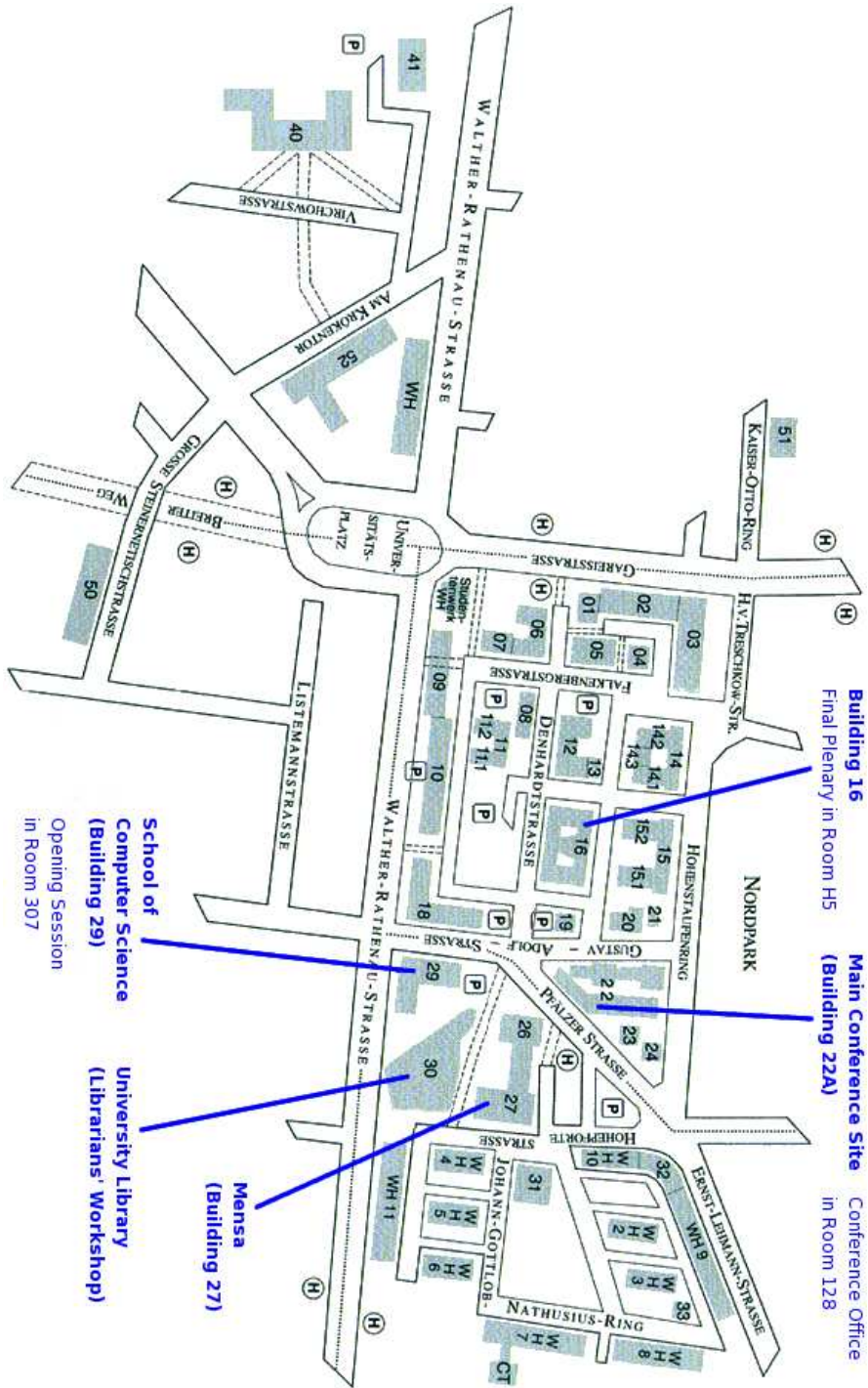
Opening Session/First Plenary Lecture:	Building 29, Room 307
Second Plenary Lecture:	Building 16, Room H5
Doctoral Consortium:	Building 29, Room 301
Librarian Workshop:	University Library

See the map on the next page for the location of the buildings. You'll find detailed travel information on the conference web pages. Please visit the conference office for registration before the opening session, we recommend to register already on tuesday evening.

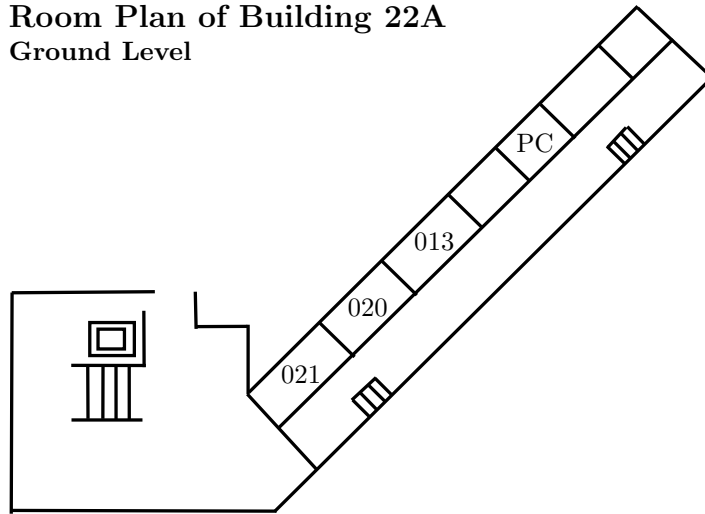
## Conference Office

Location:	Building 22A (main conference site), Room 128
Phone:	+49 391 67 18414
URL:	<a href="http://omen.cs.uni-magdeburg.de/itikmd/gfk12005/">http://omen.cs.uni-magdeburg.de/itikmd/gfk12005/</a>
Opening hours:	March 8: 17:00 - 18:00
	March 9: 08:00 - 19:00
	March 10: 08:00 - 18:00
	March 11: 08:00 - 16:00

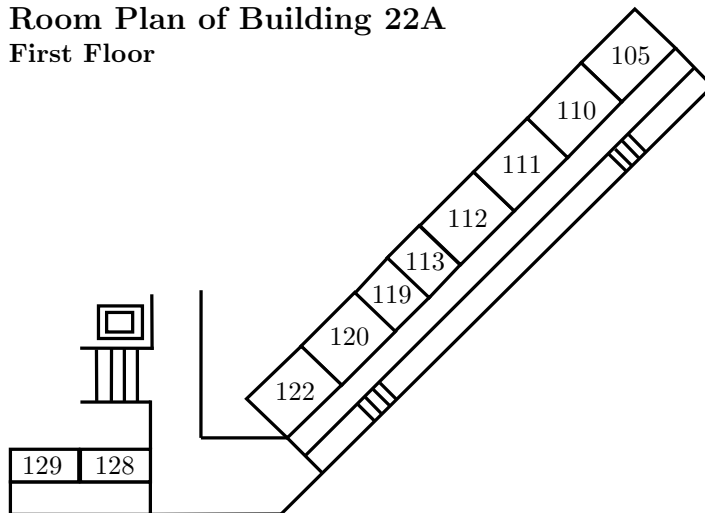
See the map on page 13 for the location of the conference office.



**Room Plan of Building 22A**  
**Ground Level**



**Room Plan of Building 22A**  
**First Floor**



The **conference office** is in Building 22A, first floor, room 128.

We provide also an **Internet room** (the “PC room”) in building 22A, ground level, room 012 (see the room plan above).

## Information for Speakers

Please establish contact with the responsible Session Chair at the beginning of the session. Beamers are located in all conference rooms. You should verify that your laptop cooperates with the beamer no later than the last break before your talk. If you need additional equipment, e.g. laptop, overhead projector or transparencies, please inform the Conference Office in the morning of the day of your talk (at the latest).

## Information for Session Chairs

The GfKl 2005 relies greatly on the involvement of the Session Chairs for a smooth flow of the presentations. The schedule is very tight, so timing is critical. Please make sure that all presentations finish within the planned time slot. To avoid delays due to equipment incompatibilities, please establish contact with all presenters in your session and make sure that they have the necessary equipment to make their presentations and that this equipment is compatible with the beamer in the room. There are auxiliary laptops, which we can schedule for the sessions if we are told early enough.

## Overview of the Registration Fees

On-site registration: March 8–11, 2005 in the conference office.

### Registration Fees (after January 31, 2005)

Regular Participant Package:

- Members of the GfKl and its associated societies,  
members of SLAIS  
(reception and proceedings included) 100,- EUR
- Non-members:  
(reception and proceedings included) 130,- EUR

Student Participant Package: 30,- EUR

Reduced Rate Package:

- Librarians and archivists participating  
in the Library Workshop 50,- EUR

## Special Events during the Conference

### General assembly of the GfKI

The general assembly of the GfKI is held on  
March 10, 2005, 17:30–19:30 hours, Building 22A, Room 021.

### Meetings of Working Groups

- AG Klassifikation und Datenanalyse in den Biowissenschaften (AG-BT)  
March 9, 2005, 12:30–13:00 hours, Building 22A, Room 105
- AG Bibliotheken (AG-BIB)  
March 9, 2005, 14:00–14:15 hours, University Library
- AG Dezimalklassifikationen (AG-DK)  
March 10, 2005, 10:00–11:15 hours, University Library
- AG Datenanalyse und Numerische Klassifikation (AG-DANK)  
March 10, 2005, 13:00–13:20 hours, Building 22A, Room 105

### Further Events

- Doctoral Workshop  
Workshop Organizer: Werner Esswein (TU Dresden)  
March 8, 2005, 13:00–17:00 hours, Building 29, Room 301  
Workshop participation is by invitation only.
- Library Workshop — Subject-Indexing 2005  
Workshop Chairs:  
Hans-J. Hermes (TU Chemnitz),  
Bernd Lorenz (Fachhochschule für Öffentliche Verwaltung  
und Rechtspflege, München)  
March 9, 2005, 10:00–17:00 hours, and  
March 10, 2005, 10:00–16:00 hours, University Library  
The Workshop is open for librarians and archivists only.
- Industrial Track
  - Industrial Application and Solutions (Prudsys)  
March 10, 2005, 9:50–11:05 hours, Building 22A, Room 120
  - Industrial Application and Solutions (SPSS)  
March 10, 2005, 15:30–17:15 hours, Building 22A, Room 111

## **Social Program**

### **Come Together**

On the evening before the conference, an informal Come Together takes place at the Restaurant Ratswaage (Ratswaageplatz 1–4) near the center of Magdeburg (see the map on page 20).

### **Reception**

The reception by the city authorities will take place on March 10, 19:00 hours, in the Hotel Ratswaage (Ratswaageplatz 1–4).

During the reception the awards will be given to the winners of the Data Mining Competition.

### **Conference Dinner**

The Conference Dinner takes place in the arch cellar (“Gewölbekeller Buttergasse”, Alter Markt 13, see the map on the next page).

Located inside a modern building, the restaurant invites you in the 800-year old cellars of medieval Magdeburg. We offer a buffet with local and vegetarian dishes and have planned a scientific-cultural event for you: we feature the music band “Musichaos”, which consists of two local professors of the Business Computer Science who will entertain you with Spanish traditional music.

Meeting point: Gewölbekeller Buttergasse, 19:30 hours  
Location: Alter Markt 13, 39104 Magdeburg  
Price: 35 Euro (beverages included)

We will meet at the restaurant. Most of the recommended hotels are located less than 10 minutes away.



- 1 — Conference Site
- 2 — Site of the Reception
- 3 — Site of the Conference Dinner

## Exploring Magdeburg

### History Written in Stone

The history of Magdeburg began with the Cathedral Square where back in time the Emperor's splendid palatinate stood. Today this place is dominated by the Magdeburg Cathedral and marvelous baroque highlights. The Town Hall is the administrative center of the capital city. In front of the building you can admire the Magdeburg Horseman—a beautifully crafted bronze cast from the first free-standing equestrian statue to be found north of the Alps, created around 1240. The original is exhibited in the Magdeburg Cultural History Museum. In order to learn more about the city's fascinating history, take a stroll along Breiter Weg and marvel at the baroque patrician houses bearing silent witness to the splendor of bygone times. Fortunately, Magdeburg is one of those cities where fortresses and ramparts tell of a rich past, such as the "Kiek in de Küken" tower (peep into the kitchen). The name was bestowed upon the tower because the wardens could peep into the kitchen of the archbishop's palace.

You have several possibilities to explore the history of Magdeburg. For more information please refer to the brochure provided in the conference bag.

### Extended City Tour

This extended tour takes you back to Magdeburg's rich past—the city is nearly 1200 years old—and includes important buildings such as the Cathedral and Monastery as well as modern sights.

Duration: 4 hours  
Time: daily excepts Mondays  
Costs: 9.00 EUR per person  
price incl. tour of Cathedral and Monastery, admission charges  
(for special exhibitions 2.00 EUR extra per person,  
extra fee for tours in a foreign language: 55.00 EUR)

### Brief City Tour

This brief tour gives you an impression of Magdeburg, a city with a long history and tradition. It also shows you the sights of a vibrant and thriving modern city that has emerged from the old Magdeburg.

Duration: 2 hours  
Time: daily excepts Mondays  
Costs: 2.00 EUR per person  
(extra fee for tours in a foreign language: 25.00 EUR)  
Remarks: no guided tours inside buildings

**City tour including Cathedral**

This tour introduces you to the most beautiful side of Magdeburg, the Cathedral City on the banks of the river Elbe. Come and marvel at the breathtaking Gothic cathedral which opens its doors and history to visitors.

Duration: 3 hours  
 Time: daily excepts Mondays  
 Costs: 6.00 EUR per person  
       price incl. tour of the Cathedral  
       (extra fee for tours in a foreign language: 40.00 EUR)

**City tour including Monastery**

This tour takes you back in history to discover the rich Romanesque heritage. The Monastery of Our Lady is not only considered the pearl but also the important centerpiece of the “Romanesque Road.”

Duration: 3 hours  
 Time: daily except Mondays  
 Costs: 6.00 EUR per person  
       price incl. Monastery tour and admission charges  
       (for special exhibitions 2.00 EUR extra per person,  
       extra fee for tours in a foreign language: 40.00 EUR)

**Tour to the Waterway Cross and Ship Lift**

Enjoy a trip to the gigantic waterway cross Magdeburg. A waterway directly connects the Mittellandkanal and the Elbe-Havel-Kanal. The centerpiece is the construction of the longest canal bridge (918m) in Europe.

Duration: 2.5 hours  
 Costs: 80.00 EUR per coach/bus  
       (extra fee for tours in a foreign language: 40.00 EUR)

**Kunstmuseum Kloster Unser Lieben Frauen**

The Monastery is one of the best preserved buildings in Romanesque style and important venue for contemporary art exhibitions in Saxony-Anhalt.

Opening hours: Tue–Sun, 10:00–17:00 hours  
 Admission: free (except special exhibitions)

**Gruson-Gewächshäuser**

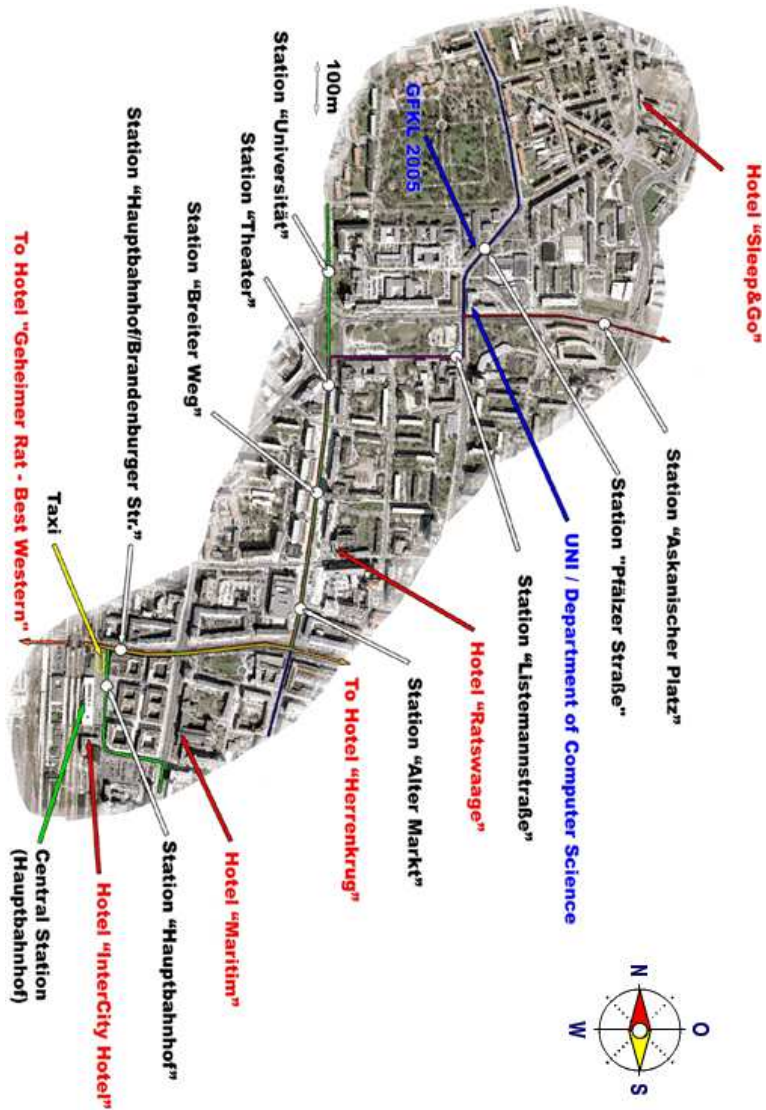
Exotic plant collection by the Magdeburg industrialist Dr. Hermann Gruson given to the city in 1895. The greenhouses include 3.000 species of tropical and subtropical plants.

Opening hours: Tue–Sun, 09:00–17:00 hours  
 Costs: Adults: 1.80 EUR  
       Concessions: 1.00 EUR (groups with 10 persons)  
       Guided tour: 2.60 EUR

## Public Transportation

Many of the hotels are at walking distance from the conference location, as can be seen on the map below.

For visiting Magdeburg, you can use public transportation: buses and trams. You can buy 1.50 EUR tickets at the conference office or at the bus/tram stations.



## Restaurants and Bistros

You can find several options for lunch near the conference location, some of which are listed below.

### **Mensa of the University**

Location: Building 27 on university campus

Distance: approximately 2 minutes to walk

Website: [www-a.uni-magdeburg.de:8080/~bischlev/speiseplan.php](http://www-a.uni-magdeburg.de:8080/~bischlev/speiseplan.php)

### **Taipan** (Thai/Chinese restaurant)

Location: Listemannstraße 18

Distance: approximately 3 minutes to walk

Website: [www.taipan-restaurant.de](http://www.taipan-restaurant.de)

### **Korfu** (Greek restaurant)

Location: Listemannstraße 18

Distance: approximately 3 minutes to walk

Website: [www.korfu-md.de.vu](http://www.korfu-md.de.vu)

### **Quartiere Latino Ristorante Gelateria** (Italian restaurant)

Location: Erzberger Straße 1

Distance: approximately 5 minutes to walk

Website: [www.ristorante-quartiere-latino.de](http://www.ristorante-quartiere-latino.de)

Several bistros can be found on Breiter Weg (pedestrian area south of the Universitätsplatz), about 10 minutes walk from the conference site.

Further restaurants, bistros, cafés and bars are around the Hasselbachplatz, close to the Magdeburg Cathedral, at one of the most picturesque locations of the city.

Further information can be found under

<http://omen.cs.uni-magdeburg.de/itikmd/index.php?id=132>



# Program



Time	Wednesday, March 9, 2005								(details: see page 29)
09:00–10:00	<b>Opening Session</b> Room 29.307 (Chair: T. Strothotte — Dean of the Faculty of Computer Science)								
10:00–10:45	Plenary Lecture W Room 29.307 <b>Support Vector Machines for Structured Outputs</b> T. Joachims (Chair: M. Spiliopoulou)								
10:45–11:15	Coffee Break								
11:15–12:30 (3 talks)	W1.1 22A.110 <b>Classification and Analysis in Data Intensive Scenarios 1</b> G. Saake	W1.2 22A.013 <b>Text Mining 1: Similarity Measures and Encoding</b> A. Nürnberger	W1.3 22A.020 <b>Clustering 1: Mixtures, Modes, and Stability</b> D. Vicari	W1.4 22A.111 <b>Classification with Latent Variable Models</b> A. Montanari	W1.5 22A.112 <b>Bioinformatics and Biostatistics 1</b> B. Lausen	W1.6 22A.120 <b>Banking and Finance 1</b> H. Locarek-Junge	W1.7 22A.105 <b>User and Data Authentication in IT Security 1</b> J. Dittmann		
12:30–14:00	Lunch Break								
	12:30–13:00: Meeting <b>Working Group “Classification and Data Analysis in Biology and Medicine”</b> BT, Room 22A.105								
14:00–14:40	Semi-Plenary W1 Room 22A.013 <b>Graph Mining: Prediction and Learning with Graph-Structured Objects</b> S. Wrobel (Chair: E. Hüllermeier)			Semi-Plenary W2 Room 22A.020 <b>Boosting and <math>L^1</math>-penalty Methods for High-dimensional Data; with some Applications in Genomics</b> P. Bühlmann (Chair: H.-H. Bock)		Semi-Plenary W3 Room 22A.021 <b>Inferred Causation, Nonparametric Methods, and Agent-Based Simulation: Candidates for a Paradigm Shift in Marketing Science?</b> J. Mazanec (Chair: O. Opitz)			
14:45–16:25 (4 talks)	W2.1 22A.013 <b>Exploiting Semantics in Web Mining</b> M. Spiliopoulou	W2.2 22A.110 <b>Text Mining 2: Linguistics 1</b> P. Grzybek	W2.3 22A.020 <b>Clustering 2: Algorithms</b> H.-H. Bock	W2.4 22A.111 <b>Classification of High-dimensional Biological and Medical Data 1</b> S. Kropf	W2.5 22A.112 <b>Bioinformatics and Biostatistics 2</b> B. Lausen	W2.6 22A.120 <b>Marketing 1</b> D. Baier / M. Meyer	W2.7 22A.105 <b>User and Data Authentication in IT Security 2</b> J. Dittmann		
16:25–16:45	Coffee Break								
16:45–18:25 (4 talks)	W3.1 22A.013 <b>Web Usage Analysis</b> A. Hotho	W3.2 22A.110 <b>Text Mining 3: Linguistics 2</b> H. Goebel	W3.3 22A.020 <b>Clustering 3: CART and Visualization</b> C. Hennig	W3.4 22A.111 <b>Classification of High-dimensional Biological and Medical Data 2</b> U. Mansmann	W3.5 22A.112 <b>Medical and Health Sciences</b> K.-D. Wernecke	W3.6 22A.120 <b>Marketing 2</b> D. Baier / M. Meyer	W3.7 22A.105 <b>Adaptivity and Personalization 1</b> A. Geyer-Schulz / L. Schmidt-Thieme		
19:00–	Reception								

**Librarians' program** 10:00–17:00 in the library: Arbeitsgruppe Bibliotheken — bibliothekarische Fortbildung (see page 53)

Timetable: Wednesday, March 9, 2005

Time	Thursday, March 10, 2005							(details: see page 38)
09:00–09:45	Semi-Plenary T1 Room 22A.013 <b>Scaling and Fortifying Fuzzy Clustering for Data Analysis</b> L.O. Hall (Chair: R. Kruse)		Semi-Plenary T2 Room 22A.020 <b>Text Mining in Action</b> D. Mladenić (Chair: A. Nürnberger)		Semi-Plenary T3 Room 22A.021 <b>Striving for an Adequate Vocabulary - Next Generation 'Metadata'</b> D.W. Fellner (Chair: A. Geyer-Schulz)			
09:50–11:05 (3 talks)	T1.1 22A.013 <b>Fuzzy Data Analysis 1</b> M. Berthold	T1.2 22A.020 <b>Text Mining 4: Clustering and Topic Detection</b> B. Stein	T1.3 22A.105 <b>Discrimination 1: Feature Selection</b> G. Ritter	T1.4 22A.110 <b>Robust Methods in Multivariate Statistics 1</b> A. Cerioli	T1.5 22A.111 <b>Data Mining Competition</b> J. Strackeljan	T1.6 22A.112 <b>Economics 1</b> O. Opitz	T1.7 22A.120 <b>Industrial Applications and Solutions (Prudsys)</b> M. Spiliopoulou	
11:05–11:20	Coffee Break							
11:20–13:00 (4 talks)	T2.1 22A.013 <b>Fuzzy Data Analysis 2</b> F. Klawonn	T2.2 22A.020 <b>Text Mining 5: Learning Ontologies</b> C. Borgelt	T2.3 22A.105 <b>Discrimination 2: PCA and Support Vectors</b> G. Ritter	T2.4 22A.110 <b>Robust Methods in Multivariate Statistics 2</b> A. Cerioli	T2.5 22A.111 <b>Classification and Analysis in Data Intensive Scenarios 2</b> S. Conrad	T2.6 22A.112 <b>Marketing 3</b> D. Baier / M. Meyer	T2.7 22A.120 <b>Adaptivity and Personalization 2</b> A. Geyer-Schulz / L. Schmidt-Thieme	
13:00–14:15	Lunch Break							
	13:00–13:25: Meeting <b>Working Group DANK</b> , Room 22A.105							
14:15–15:00	Semi-Plenary T4 Room 22A.013 <b>Multimedia Mining: The Text Is Not Enough</b> M. Detyniecki (Chair: C. Borgelt)		Semi-Plenary T5 Room 22A.020 <b>From Technology driven Research to Hypothesis based Questioning: Microarrays in Molecular Medicine</b> U. Mansmann (Chair: C. Weihs)		Semi-Plenary T6 Room 22A.021 <b>Identification of Real-World Objects in Multiple Databases</b> M. Neiling (Chair: H.-J. Lenz)			
15:00–15:30	Coffee Break							
15:30–17:15 (4 talks)	T3.1 22A.013 <b>Fuzzy Data Analysis 3</b> R. Mikut	T3.2 22A.020 <b>Text Mining 6: Information and Relation Extraction</b> A. Hotho	T3.3 22A.105 <b>Discrimination 3: Applications of Discriminant Analysis</b> G. Ritter	T3.4 22A.110 <b>Multimode Clustering and Dimensionality Reduction</b> M. Vichi	T3.5 22A.112 <b>Marketing 4</b> D. Baier / M. Meyer	T3.6 22A.120 <b>Banking and Finance 2</b> H. Locarek-Junge	T3.7 22A.111 <b>Industrial Applications and Solutions (SPSS)</b> M. Spiliopoulou	
17:30–19:00	General Assembly of the GfKI Room: 22A.021							
20:00–	Conference Dinner							

**Posters of the Doctoral Workshop:** on display in the foyer of building 22A from 15:00 to 17:30 hours

**Librarians' program** 10:00–16:00 in the library: Bibliothekarische Sektion (see page 54)

Time	Friday, March 11, 2005 <span style="float: right;">(details: see page 46)</span>							
09:00–10:40 (4 talks)	F1.1 22A.013 <b>Mining in Business Processes</b> C. Rautenstrauch	F1.2 22A.020 <b>Text Mining 7: Classification 1</b> D. Mladenić	F1.3 22A.105 <b>Multiway Classification and Data Analysis</b> S. Krolak-Schwerdt / H.A.L. Kiers	F1.4 22A.110 <b>Computational Advances in Data Analysis</b> H.-J. Lenz	F1.5 22A.111 <b>Visualization</b> P.J.F. Groenen	F1.6 22A.112 <b>Economics 2</b> O. Opitz	F1.7 22A.120 <b>Music Analysis 1</b> C. Weihs	
10:40–11:00	Coffee Break							
11:00–12:15 (3 talks)	F2.1 22A.013 <b>Data Mining and Explorative Multivariate Data Analysis 1</b> L. D'Ambra / P. Giudici	F2.2 22A.020 <b>Text Mining 8: Classification 2</b> G. Paaß	F2.3 22A.105 <b>Ranking, Multi-Label Classification, Preferences 1</b> J. Fürnkranz / E. Hüllermeier	F2.4 22A.110 <b>Dissimilarities and Clustering Structures 1</b> B. Fichet	F2.5 22A.111 <b>PLS Path Modeling, PLS Regression and Classification 1</b> N.C. Lauro / V.E. Vinzi	F2.6 22A.112 <b>Economics 3</b> O. Opitz	F2.7 22A.120 <b>Music Analysis 2</b> F. Mörchen	
12:15–13:30	Lunch Break							
13:30–14:45 (3 talks)	F3.1 22A.013 <b>Data Mining and Explorative Multivariate Data Analysis 2</b> L. D'Ambra / P. Giudici	F3.2 22A.020 <b>Text Mining 9: Retrieval Systems</b> W. Gaul	F3.3 22A.105 <b>Ranking, Multi-Label Classification, Preferences 2</b> J. Fürnkranz / E. Hüllermeier	F3.4 22A.110 <b>Dissimilarities and Clustering Structures 2</b> B. Brucker	F3.5 22A.111 <b>PLS Path Modeling, PLS Regression and Classification 2</b> N.C. Lauro / V.E. Vinzi	F3.6 22A.112 <b>Economics 4</b> O. Opitz	F3.7 22A.120 <b>Music Analysis 3</b> D. Müllensiefen	
14:50–15:35	Plenary Lecture F Room 16.H5 <b>SolEuNet: Selected Data Mining Techniques and Applications</b> N. Lavrač (Chair: W. Gaul)							
15:35–	Informal Closing Session							

Timetable: Friday, March 11, 2005

## Doctoral Workshop

**Tuesday, March 8, 2005**

Room: 29.301

	Introduction	Werner Esswein Technische Universität Dresden
13:00	Statistische Verfahren zur Erkennung mathematischer Begabung	Meike Gebel Universität Dortmund
13:30	Klassifikation interorganisationaler Informationssysteme	Astrid Zenke RWTH Aachen
14:00	Extensions of the Clustering of Variables around Latent Components	Karin Sahmer ENITIAA/INRA, Nantes, France
14:30	coffee break	
15:00	Klassifikation von Magnetresonanz-Spektroskopischen Bildern in der Gradierung und Lokalisation von Hirntumoren	Bjoern H. Menze Ruprecht-Karls-Universität Heidelberg
15:30	Zur Optimierung bei netzeffekt-basierten Diffusionsmodellen	Karl-Heinz Lüke Universität Hildesheim
16:00	Feedback on the research results	

**Posters of the Doctoral Workshop:**

on display in the foyer of building 22A

Thursday, March 10, 2005, 15:00 to 17:30 hours

---

**Wednesday, March 9, 2005**


---

<b>Opening Session</b>	Room 29.307, 09:00–10:00
Chair: T. Strothotte — Dean of the Faculty of Computer Science	
Classification Song	
Opening Addresses	09:05–09:20
<i>President of the University of Magdeburg</i>	
<i>President of the GfKI</i>	
<i>Local Conference Chair</i>	
Best Paper Awards of GfKI 2004	09:20–09:40
Data Mining Competition Results	09:40–10:00

---

Plenary W	Room 29.307, 10:00–10:45
<b>Support Vector Machines for Structured Outputs</b>	
Speaker: T. Joachims	
Chair: M. Spiliopoulou	55

---

Session W1.1	Room 22A.110, 11:15–12:30
<b>Classification and Analysis in Data Intensive Scenarios 1</b>	
Chair: G. Saake	
Improving Integration Quality	56
<i>E. Altareva, S. Conrad</i>	
Automated Topic Labeling at the Sentence Level: Potential and Limitations	57
<i>K. Winkler, H. Graubitz, M. Spiliopoulou</i>	
From Web-Publication to the Database — Full Text Mining Upon the German Register of Trade	58
<i>L. Wille, K. Winkler, M. Spiliopoulou</i>	

---

Session W1.2	Room 22A.013, 11:15–12:30
<b>Text Mining 1: Similarity Measures and Encoding</b>	
Chair: A. Nürnberger	
Weighting Systems for Text Mining	59
<i>S. Balbi, M. Misuraca</i>	
Semantic Decomposition of Character Encodings for Linguistic Knowledge Discovery	60
<i>D. Gibbon, B. Hughes, T. Trippel</i>	
Fuzzy-Fingerprints For Near Similarity Search	61
<i>B. Stein, S. Meyer zu Eissen</i>	

---

---

Session W1.3	Room 22A.020, 11:15–12:30
<b>Clustering 1: Mixtures, Modes, and Stability</b>	
Chair: D. Vicari	
An Indicator for the Number of Modes in a Mixture Model Using a Linear Map to Simplex Structure	62
<i>M. Weber, W. Rungtarityotin, A. Schliep</i>	
Model Selection Criteria for Latent Class Models — Results from a Monte Carlo Study	63
<i>J.G. Dias</i>	
Finding Meaningful and Stable Clusters by Local Cluster Analysis	64
<i>A. Mucha</i>	

---

Session W1.4	Room 22A.111, 11:15–12:30
<b>Classification with Latent Variable Models</b>	
Chair: A. Montanari	
Identifying Multiple Cluster Structures through Latent Class Models	65
<i>G. Galimberti, G. Soffritti</i>	
Dimensionality Reduction for Discrimination by Independent Factor Analysis	66
<i>D.G. Calò, A. Montanari, C. Viroli</i>	
Variable Selection in Classification Problems via Projections onto a Latent Space	67
<i>M. Pillati, C. Viroli</i>	

---

Session W1.5	Room 22A.112, 11:15–12:30
<b>Bioinformatics and Biostatistics 1</b>	
Chair: B. Lausen	
Tests for Multiple Change Points in Binary Markov Sequences	68
<i>J. Krauth</i>	
Clustering Surface Sections of Protein Binding Cavities	69
<i>K. Kupas, G. Klebe, A. Ultsch</i>	
On External Indices for Mixtures	70
<i>I.G. Costa, A. Schliep</i>	

---

---

Session W1.6	Room 22A.120, 11:15–12:30	
<b>Banking and Finance 1</b>		
Chair: H. Locarek-Junge		
The Classification of Candlestick Charts: Laying the Foundation for Further Empirical Research <i>S. Etschberger, H. Fock, C. Klein, B. Zwergel</i>		71
Value at Risk Using Principal Components Analysis on Polish Power Exchange <i>G. Trzpiot, A. Ganczarek</i>		72
Comparing and Selecting SVM-Kernels for Credit Scoring <i>R. Steeking, K.B. Schebesch</i>		73

---

Session W1.7	Room 22A.105, 11:15–12:30	
<b>User and Data Authentication in IT Security 1</b>		
Chair: J. Dittmann		
Unimodal and Multimodal Biometric Authentication Systems based on Hand Features <i>N. Pavesic, S. Ribaric</i>		74
Hand-Geometry Recognition Based on Contour Landmarks <i>R. Veldhuis, A. Bazen, W. Booij, A. Hendrikse</i>		75
A Cross-cultural Evaluation Framework for Behavioral Biometric User Authentication <i>T.K. Basu, P.K. Dutta, C. Vielhauer, A. Oermann, B. Yegnanarayana</i>		76

---

Semi-Plenary W1	Room 22A.013, 14:00–14:40	
<b>Graph Mining: Prediction and Learning with Graph-Structured Objects</b>		
Speaker: S. Wrobel		
Chair: E. Hüllermeier		77

---

Semi-Plenary W2	Room 22A.020, 14:00–14:40	
<b>Boosting and <math>\mathcal{L}^1</math>-penalty Methods for High-dimensional Data; with some Applications in Genomics</b>		
Speaker: P. Bühlmann		
Chair: H.-H. Bock		78

---

---

Semi-Plenary W3 Room 22A.021, 14:00–14:40  
**Inferred Causation, Nonparametric Methods, and Agent-Based Simulation: Candidates for a Paradigm Shift in Marketing Science?**  
 Speaker: J. Mazanec  
 Chair: O. Opitz 79

---

Session W2.1 Room 22A.013, 14:45–16:25  
**Exploiting Semantics in Web Mining**  
 Chair: M. Spiliopoulou

Ontology-based Online Categorization of Web Pages 80  
*E.W. de Luca, A. Nürnberger*

Reconstructing Hypertext Categorization 81  
*A. Mehler, R. Gleim, M. Dehmer*

Using Formal Concept Analysis for Mining Semantic Log Files 82  
*A. Hotho, C. Schmitz, G. Stumme, J. Tane*

Determining Senses for Word Sense Disambiguation in Turkish 83  
*Z. Orhan, Z. Altan*

---

Session W2.2 Room 22A.110, 14:45–16:25  
**Text Mining 2: Linguistics 1**  
 Chair: P. Grzybek

Chronological Analysis of Corpus Data — 84  
 On the Example of the Wrocław Corpus of Polish  
*A. Pawlowski*

Reconstruction of Two-dimensional Family Trees and 85  
 Problems of Genetic Language Classification  
*A. Polikarpov, V. Kromer*

Part-of-speech Induction by Clustering Local Contexts 86  
*R. Rapp*

---

Session W2.3 Room 22A.020, 14:45–16:25  
**Clustering 2: Algorithms**  
 Chair: H.-H. Bock

On the Aggregation of Structures — 87  
*k*-Means for Data Matrices with Mixed Information Levels  
*J. Hansohm*

Robust Clustering for Nominal and Continuous Data 88  
*S. Rüping*

	Detailed Program: Wednesday, March 9, 2005	33
Classification with Numerical Taxonomy		89
<i>T. Postelnicu</i>		
U*C: Self-organized Clustering with Emergent Feature Maps		90
<i>A. Ultsch</i>		
<hr/>		
Session W2.4	Room 22A.111, 14:45–16:25	
<b>Classification of High-dimensional Biological and Medical Data 1</b>		
Chair: S. Kropf		
Reproducible Research in High-dimensional Classification Tasks		91
<i>U. Mansmann, M. Ruschhaupt, W. Huber</i>		
Search for Relevant Sets of Variables, Applied in Gene Expression Analysis		92
<i>J. Lauter</i>		
On Variable Selection for Discrimination of More Than Two Classes if the Number of Observations is Moderate		93
<i>G. Szepannek, C. Weihs</i>		
Structures of the Screening Procedures for Malnourished Patients with Gastric Cancer		94
<i>M. Gorkiewicz, E. Walewska, L. Scislo, A. Czupryna, S. Klek, A.M. Szczepanik, J. Kulig</i>		
<hr/>		
Session W2.5	Room 22A.112, 14:45–16:25	
<b>Bioinformatics and Biostatistics 2</b>		
Chair: B. Lausen		
Error reducing Aggregation of c-DNA Microarray Experiments		95
<i>A. Ultsch</i>		
Detecting Chromosomal Aberrations using Hidden Markov Models with Inhomogeneous Markov Chains		96
<i>A. Schliep, S. Roepcke</i>		
Swarms of Artificial Life Forms for Clustering DNA Microarray Data		97
<i>D. Kampf, A. Ultsch</i>		
<hr/>		
Session W2.6	Room 22A.120, 14:45–16:25	
<b>Marketing 1</b>		
Chair: D. Baier / M. Meyer		
Entropy-Based Approach to External Spatial and Non-Spatial Market Structure Analysis		98
<i>A. Sagan, M. Lapczyński</i>		

The Importance of E-Commerce in China and Russia — An Empirical Comparison <i>F. Kroll, A. Hermelbracht</i>	99
Evaluating Academic Library Services by Means of Web-based Conjoint Analysis <i>A. Hermelbracht, R. Decker</i>	100
Optimizing Profit for Bundle Sales Using Estimated Reservation Prices <i>C. Breidert, M. Hahsler</i>	101

---

Session W2.7 Room 22A.105, 14:45–16:25  
**User and Data Authentication in IT Security 2**  
 Chair: J. Dittmann

Towards a Flexible Framework for Open Source Software for Handwritten Signature Analysis <i>M. Fairhurst, R. Guest, C. Vielhauer</i>	102
Biometric User Authentication — Technical and Legal Implication for the Information Management <i>J. Hansen, C. Vielhauer</i>	103
Lossless Data Hiding for Recovering the Original Data <i>R. Caldelli, A. De Rosa, A. Piva</i>	104
Heuristics for Optimizing Global Spam-Classifying-Processes <i>W. Burkart, S. Etschberger, C. Klein</i>	105

---

Session W3.1 Room 22A.013, 16:45–18:25  
**Web Usage Analysis**  
 Chair: A. Hotho

Evaluation for Web Mining Applications <i>B. Berendt, E. Menasalvas, M. Spiliopoulou</i>	106
Classification and Rule Discovery for Temporal Web Mining <i>M. Samia, S. Conrad</i>	107
A Classification of Online Music Consumers <i>H.H. Bauer, T.E. Haber, F. Huber, M.M. Neumann</i>	108
The Impact of Experimentation in Web Mining <i>E. Menasalvas, P. Herrero, J. Segovia</i>	109

---

---

Session W3.2	Room 22A.110, 16:45–18:25
<b>Text Mining 3: Linguistics 2</b>	
Chair: H. Goebel	
Text Classification: The Impact of Sentence Length	110
<i>P. Grzybek, E. Stadlober</i>	
Word Length and Frequency Distributions	111
<i>G. Antić, E. Kelih</i>	
Crosslinguistic Computation and a Rhythm-based Classification	112
<i>G. Fenk-Oczlon, A. Fenk</i>	
Proposals for the Enhancement of Unsupervised Segmentation of Words using Successor Values	113
<i>C. Benden</i>	

---

Session W3.3	Room 22A.020, 16:45–18:25
<b>Clustering 3: CART and Visualization</b>	
Chair: C. Hennig	
I2C4.5: An Innovative Discretisation Technique for C4.5, Using Intense Intervals	114
<i>I. Bourlacos, A. Sotiropoulou, D. Theotokis</i>	
PRISMA: Improving Risk Estimation with Parallel Logistic Regression Trees	115
<i>B. Arnrich, A. Albert, J. Walter</i>	
Visualization Techniques for Semi-Automatic Subgroup Mining	116
<i>M. Atzmueller, T. Vogeles, F. Puppe</i>	
A Strategy for Visualizing and Classifying Keywords and Units in Textual Data Analysis	117
<i>M.G. Grassia, M. Misuraca, G. Scepi</i>	

---

Session W3.4	Room 22A.111, 16:45–18:25
<b>Classification of High-dimensional Biological and Medical Data 2</b>	
Chair: U. Mansmann	
Predicting Survival using Microarray Gene Expression Data	118
<i>A. Benner, C. Itrich, U. Mansmann</i>	
Methods for Class Discovery in High-dimensional and Noisy Gene Expression Data based on Data Resampling and Reclustering	119
<i>U. Möller, D. Radke</i>	
A Ridge Classification Method for High-dimensional Observations	120
<i>M. Grüning, S. Kropf</i>	

Knowledge Discovery for Grape Musts to Optimize the Production of Wine 121  
*A. Ulsch, F. Mörchen, C.-D. Patz, M. Hieber*

Session W3.5 Room 22A.112, 16:45–18:25

**Medical and Health Sciences**

Chair: K.-D. Wernecke

Predictive Accuracy of some Indices used to Screen the Malnourished Patients with Aim to Facilitate Targeted Nutritional Intervention 122  
*E. Walewska, L. Scislo, M. Górkiewicz, A. Czupryna, S. Klek, A.M. Szczepanik, J. Kulig.*

Heart Rate Classification Using Support Vector Machines 123  
*M. Vogt, U. Moissl, J. Schaab*

Methods and Pitfalls in the Assessment of Second and Secondary Cancers 124  
*S. De Cantis, A.M. Taormina, E. Iannitto*

Dichotomizing Continuous Prognostic Variables: When, Why and How? 125  
*S. De Cantis, A.M. Taormina, E. Iannitto*

Session W3.6 Room 22A.120, 16:45–18:25

**Marketing 2**

Chair: D. Baier / M. Meyer

Stability of Consumers' Opinions — Measurement and Classification Attempt 126  
*S. Jadwiga*

Applying Support Vector Machines in Marketing 127  
*G. Nalbantov, J.C. Bioch, P.J.F. Groenen*

Modelling the Non-linear Relationship Between Satisfaction and Loyalty with Structural Equation Models 128  
*M. Paulssen, A. Sommerfeld*

An Exploratory Analysis of the Relationship Between Customer's and Agent's Product Affinity 129  
*M. Heede, M. Meyer*

Session W3.7 Room 22A.105, 16:45–18:25

**Adaptivity and Personalization 1**

Chair: A. Geyer-Schulz / L. Schmidt-Thieme

About the Use of "Intelligent" Search Engines for the Exchange of Environmental Data 130  
*E. Weihs*

Detailed Program: Wednesday, March 9, 2005 37

Attribute-aware Collaborative Filtering	131
<i>K. Tso, L. Schmidt-Thieme</i>	
Forum Based Collaborative Knowledge Creation and Knowledge Sharing: Content Based Knowledge Process Derivation	132
<i>F. Schmitz-Justen, A.F.X. Wilhelm</i>	
User Modeling for a Scientific Information Portal	133
<i>E. Schwarzkopf</i>	

---

---

**Thursday, March 10, 2005**


---

Semi-Plenary T1 Room 22A.013, 09:00–09:45  
**Scaling and Fortifying Fuzzy Clustering for Data Analysis**  
 Speaker: L.O. Hall  
 Chair: R. Kruse 134

---

Semi-Plenary T2 Room 22A.020, 09:00–09:45  
**Text Mining in Action**  
 Speaker: D. Mladenić  
 Chair: A. Nürnberger 135

---

Semi-Plenary T3 Room 22A.021, 09:00–09:45  
**Striving for an Adequate Vocabulary - Next Generation ‘Metadata’**  
 Speaker: D.W. Fellner  
 Chair: A. Geyer-Schulz 136

---

Session T1.1 Room 22A.013, 09:50–11:05  
**Fuzzy Data Analysis 1**  
 Chair: M. Berthold

Fuzzy, Neural, and Statistical Classifiers in Medical Data Mining — 137  
 Problem Formulation, Performance and Interpretability  
*R. Mikut, M. Reischl, T. Loose*

Classification of Ultrasonic Signals for the Detection of Tooth Surfaces 138  
*J. Strackeljan, B. Oehme*

Autonomous Sensor Based Landing Systems: Fusion of Vague and In- 139  
 complete Information by Application of Fuzzy Clustering Techniques  
*B. Korn*

---

Session T1.2 Room 22A.020, 09:50–11:05  
**Text Mining 4: Clustering and Topic Detection**  
 Chair: B. Stein

Transcript Topic Tracking and Segmentation: 140  
 A New Application for Lexical Chains  
*C. Chibelushi, B. Sharp, J. Di Manno*

Applying Collaborative Filtering to Real-life Corporate Data 141  
*M. Grcar, D. Mladenić, M. Grobelnik*

Multilingual Text Corpus Classification Using Thesaurus EuroWordNet 142  
*M. Toman, K. Ječek*

---

---

Session T1.3	Room 22A.105, 09:50–11:05
<b>Discrimination 1: Feature Selection</b>	
Chair: G. Ritter	
Input Variable Selection in Kernel Fisher Discriminant Analysis	143
<i>N. Lowy, S. Steel</i>	
A Wrapper Feature Selection Method for Combined Tree-based Classifiers	144
<i>E. Gatnar</i>	
Countering the False Positive Projection Effect in Nonlinear Asymmetric Classification	145
<i>S. Kosinov, S. Marchand-Maillet, T. Pun</i>	

---

Session T1.4	Room 22A.110, 09:50–11:05
<b>Robust Methods in Multivariate Statistics 1</b>	
Chair: A. Cerioli	
Deepest Points and Least Deep Points: Robustness and Outliers with MZE	146
<i>C. Becker</i>	
Robust Multivariate Methods: The Projection Pursuit Approach	147
<i>P. Filzmoser</i>	
Robust Transformations and Outlier Detection with Correlated Data	148
<i>M. Riani, A. Cerioli</i>	

---

Session T1.5	Room 22A.111, 09:50–11:05
<b>Data Mining Competition</b>	
Chair: J. Strackeljan	
Predicting Liquidity Crisis in Companies ( <i>presented in opening session</i> )	149
<i>D.S. Vogel</i>	
Predicting Liquidity Crises of Companies — Part I: Data Preprocessing	150
<i>I. Bezrukov, T. Deselaers, A. Hegerath, D. Keysers, A. Mauser</i>	
Predicting Liquidity Crises of Companies — Part II: Training and Classification	151
<i>A. Mauser, D. Keysers, A. Hegerath, T. Deselaers, I. Bezrukov</i>	

---

---

Session T1.6	Room 22A.112, 09:50–11:05
<b>Economics 1</b>	
Chair: O. Opitz	
Regional GDP and Employment Forecasting	152
<i>W. Polasek, H. Berrer</i>	
An Analysis Method for the Structure of a Political Stock Market and its Efficiency	153
<i>M. Franke, A. Geyer-Schulz, B. Hoser</i>	
Influence of Information on Individual Decision Behavior	154
<i>J. Kunze</i>	

---

Session T1.7	Room 22A.120, 09:50–11:05
<b>Industrial Applications and Solutions (Prudsys)</b>	
Chair: M. Spiliopoulou	
Multilevel Approximation Theory for the Solution of Large Regression Problems	155
<i>M. Thess</i>	
Optimisation and Improvement of a Recommendation Engine	156
<i>S. Rost</i>	
Data Preparation for Data Mining in Mail Order Business	157
<i>H. Sieber</i>	

---

Session T2.1	Room 22A.013, 11:20–13:00
<b>Fuzzy Data Analysis 2</b>	
Chair: F. Klawonn	
Benefitting from Missing Values for Fuzzy Rule Formation	158
<i>T.R. Gabriel, M.R. Berthold</i>	
Objective Function-Based Classification	159
<i>F. Höppner</i>	
Outlier Preserving Clustering for Structured Data through Kernels	160
<i>M.-J. Lesot, B. Bouchon-Meunier</i>	
The Fitting of Multifunctions: An Approach to Nonparametric Multimodal Regression	161
<i>J. Einbeck, G. Tutz</i>	

---

---

Session T2.2	Room 22A.020, 11:20–13:00
<b>Text Mining 5: Learning Ontologies</b>	
Chair: C. Borgelt	
Learning Ontologies to Improve Text Clustering and Classification	162
<i>S. Bloehdorn, P. Cimiano, A. Hotho</i>	
Automatic Acquisition of Lexical and Ontological Knowledge	163
<i>U. Rist</i>	
Automatic Extension of Feature-based Semantic Lexicons via Contextual Features	164
<i>C. Biemann, R. Osswald</i>	
Clustering for Ontology Evolution	165
<i>R. Pitkanen, G. Tsatsaronis, M. Vazirgiannis</i>	

---

Session T2.3	Room 22A.105, 11:20–13:00
<b>Discrimination 2: PCA and Support Vectors</b>	
Chair: G. Ritter	
New Effective Algorithm for Stepwise Principle Components Selection in Discriminant Analysis	166
<i>E. Serikova, E. Zhuk</i>	
Generalized Relevance Learning Vector Quantization with Local Metric Adaptation in Classification of Mass Spectrometric Data	167
<i>F.-M. Schleif, T. Villmann, B. Hammer, U. Clauss</i>	
A Comparison of Validation Methods for Learning Vector Quantization and for Support Vector Machines on Two Biomedical Data Sets	168
<i>D. Sommer, M. Golz</i>	

---

Session T2.4	Room 22A.110, 11:20–13:00
<b>Robust Methods in Multivariate Statistics 2</b>	
Chair: A. Cerioli	
Hierarchical Clustering by Means of Model Grouping	169
<i>C. Agostinelli, P. Pellizzari</i>	
Robustness Concepts for General Clustering Methods	170
<i>C. Hennig</i>	
On Mean Estimation and Population Stratification	171
<i>M. Skibicki</i>	
Finding Persisting States in Univariate Time Series	172
<i>F. Mörchen, A. Ultsch</i>	

---

---

Session T2.5	Room 22A.111, 11:20–13:00
<b>Classification and Analysis in Data Intensive Scenarios 2</b>	
Chair: S. Conrad	
Distributed Clustering on Dynamic Data Sets in P2P Networks	173
<i>M. Eisenhardt, W. Müller, A. Henrich</i>	
Processing Sequential Patterns in Relational Databases	174
<i>X. Shang, E. Schallehn, K.-U. Sattler</i>	
Wrapper Generation for Meta-Search	175
<i>B. Stein, S. Meyer zu Eissen</i>	
CAMbase — XML-based Mining, Processing and Classification of Bibliographical Data on Complementary and Alternative Medicine (CAM)	176
<i>T. Ostermann, H. Zillmann, P.F. Matthiessen</i>	

---

Session T2.6	Room 22A.112, 11:20–13:00
<b>Marketing 3</b>	
Chair: D. Baier / M. Meyer	
Construction of Consensus Partitions Based on Vector Quantisation of Market Basket Data	177
<i>T. Reutterer, K. Hornik, M. Hahsler</i>	
Can Diffuse Numerical Responses be Analyzed Using Standard Techniques?	178
<i>B. Vogt, T. Spengler</i>	
Determination of the Derived Importance of Employers' Characteristics during Job Choice using Logistic Regression	179
<i>T. Spengler, J. Malmendier</i>	
Semiparametric Stepwise Regression to Estimate Sales Promotion Effects	180
<i>W. Steiner, C. Belitz, S. Lang</i>	

---

Session T2.7	Room 22A.120, 11:20–13:00
<b>Adaptivity and Personalization 2</b>	
Chair: A. Geyer-Schulz / L. Schmidt-Thieme	
Implications of Probabilistic Data Modeling for Mining Associations	181
<i>M. Hahsler, K. Hornik, T. Reutterer</i>	
Hierarchical Authorization for Personalized Access to a Web Based Digital Library	182
<i>A. Geyer-Schulz, A. Thede</i>	

Copula Functions in Model Based Clustering <i>K. Jajuga, D. Papla</i>	183
Experimenting with Document Representation <i>A. Gorur, M. Tolun, H. Sever</i>	184
<hr/>	
Semi-Plenary T4 <b>Multimedia Mining: The Text Is Not Enough</b> Speaker: M. Detyniecki Chair: C. Borgelt	Room 22A.013, 14:15–15:00 185
<hr/>	
Semi-Plenary T5 <b>From Technology driven Research to Hypothesis based Questioning: Microarrays in Molecular Medicine</b> Speaker: U. Mansmann Chair: C. Weihs	Room 22A.020, 14:15–15:00 186
<hr/>	
Semi-Plenary T6 <b>Identification of Real-World Objects in Multiple Databases</b> Speaker: M. Neiling Chair: H.-J. Lenz	Room 22A.021, 14:15–15:00 187
<hr/>	
Session T3.1 <b>Fuzzy Data Analysis 3</b> Chair: R. Mikut	Room 22A.013, 15:30–17:15
Improved Surface Defect Classification for Quality Control of Car Body Panels <i>C. Döring, A. Eichhorn, R. Kruse</i>	188
Semi-Supervised Learning of Fuzzy Classifiers <i>A. Klose</i>	189
Application of Data Mining Methods for the Prediction of Delay at Frankfurt Airport <i>F. Rehm, F. Klawonn</i>	190
A Generalised Flexible Fuzzy Clustering Approach to Avoid Highly Ambiguous Membership Degrees <i>F. Klawonn</i>	191
<hr/>	
Session T3.2 <b>Text Mining 6: Information and Relation Extraction</b> Chair: A. Hotho	Room 22A.020, 15:30–17:15
Discovering Communities in Linked Data by Multi-View Clustering <i>I. Drost, S. Bickel, T. Scheffer</i>	192

Analysis of the Authors Collaboration Networks	193
<i>V. Batagelj, S. Korenjak-Černe, M. Zaveršnik</i>	
Transitive Text Mining for Information Extraction and Hypothesis Generation	194
<i>J. Stegmann, G. Grohmann</i>	
A Hybrid Machine Learning Approach to Information Extraction	195
<i>G. Neumann</i>	

---

Session T3.3 Room 22A.105, 15:30–17:15

**Discrimination 3: Applications of Discriminant Analysis**

Chair: G. Ritter

On the Use of Some Classification Quality Measure to Construct Mean Value Estimates under Nonresponse	196
<i>W. Gamrot</i>	
The Wavelet Packet Based Cepstral Features for Open Set Speaker Classification in Marathi	197
<i>H.A. Patil, P.K. Dutta, T.K. Basu</i>	
Prediction Optimal Classification of Business Phases	198
<i>K. Luebke, C. Weihs</i>	
Discriminant Analysis of Polythetically Described Older Palaeolithic Flakes: Possibilities and Questions	199
<i>T. Weber</i>	

---

Session T3.4 Room 22A.110, 15:30–17:15

**Multimode Clustering and Dimensionality Reduction**

Chair: M. Vichi

Interpretable Clustering via Model-Based Divisive Hierarchical Classification	200
<i>N. Sujka, G. Govaert, C. Ambroise</i>	
Global Optimization Strategies for Two-mode Clustering	201
<i>J. van Rosmalen, P.J.F. Groenen, J. Trejos, W. Castillo</i>	
Multimode Partitioning	202
<i>R. Rocci, M. Vichi</i>	
Multivariate Prediction Including Clustering and Dimensionality Reduction	203
<i>D. Vicari, M. Vichi</i>	

---

---

Session T3.5	Room 22A.112, 15:30–17:15
<b>Marketing 4</b>	
Chair: D. Baier / M. Meyer	
Bayesian Versus Traditional Procedures for Latent Class Metric Conjoint Analysis	204
<i>D. Baier</i>	
Computer-Assisted Decision Tools for the Support of the Innovation Process	205
<i>M. Queitsch, D. Baier</i>	
A Market Basket Analysis Based on the Multivariate MNL	206
<i>Y. Boztug, L. Hildebrandt</i>	
Scenario Technique for Organizational Visioning in the Automotive	207
<i>T. Teichert, H. Hetzer, J.-P. Kuhle</i>	

---

Session T3.6	Room 22A.120, 15:30–17:15
<b>Banking and Finance 2</b>	
Chair: H. Locarek-Junge	
Charting Patterns Recognition and Classification on Price History	208
<i>S. Staleuskaya</i>	
Credit Portfolio Loss Distribution and Conditioning	209
<i>S. Höse, K. Vogl</i>	
Data Snooping, Multiple Hypothesis Testing and the Performance of Technical Trading Rules	210
<i>M. Papatrifon</i>	
Dynamical Systems in Capital Accumulation Models and Development Planning	211
<i>C. Majumdar</i>	

---

Session T3.7	Room 22A.111, 15:30–17:15
<b>Industrial Applications and Solutions (SPSS)</b>	
Chair: M. Spiliopoulou	
Combining Text Mining and Data Mining for Improved Business Benefit	212
<i>T. Khabaza, E.P.G. Martin, O. Jouve</i>	
Using Decision Trees To Classify Customer Value Segments	213
<i>S. Grube</i>	
Models of Web Navigation in e-Commerce	214
<i>M. Hudec, R. Kurzawa, A. Hacker</i>	

---

---

**Friday, March 11, 2005**


---

Session F1.1 Room 22A.013, 09:00–10:40

**Mining in Business Processes**

Chair: C. Rautenstrauch

Data Mining Based Diagnosis in Resource Management 215

*M. Beck, J. Marx-Gómez*

Intrusion Detection Management System for eCommerce Security 216

*J. Lichtenberg, J. Marx-Gómez*

Automated Classification of Large Product Databases 217

According to Industrial Standards

*M. Romacker, K. Marko, M. Bartel*

An Overview of Artificial Life Approaches for Clustering 218

*D. Kämpf, A. Ultsch*

---

Session F1.2 Room 22A.020, 09:00–10:40

**Text Mining 7: Classification 1**

Chair: D. Mladenić

A Framework for Conceptual Document Classification 219  
based on an Ontology of Proper Names

*D. Vitas, G. Pavlovic-Lazetic, C. Krstev*

Automatic Discovery and Classification of Web Page Section Categories 220

*S.R. El-Beltagy, M. Hazman, A. Rafea, S. ElGamal*

Classification of Documents into Hierarchy Using String Kernels 221

*B. Fortuna, M. Grobelnik, D. Mladenić*

Text Classification in the Area of Personal Recommendations 222

*C. Bomhardt, W. Gaul*

---

Session F1.3 Room 22A.105, 09:00–10:40

**Multiway Classification and Data Analysis**

Chair: S. Krolak-Schwerdt / H.A.L. Kiers

Probabilistic Models for Subjective Multiway Classifications 223  
of Memories

*T. Meiser*

Hierarchical Classes Analysis for Real-valued Data 224

*J. Schepers, I. Van Mechelen*

Selecting among Three-way Principal Component Models of Different Types and Complexities: A Numerical Convex Hull based Method <i>E. Ceulemans, H.A.L. Kiers</i>	225
---	-----

The Recovery Performance of Two-mode Clustering Methods: A Monte Carlo Experiment <i>S. Krolak-Schwerdt, M. Wiedenbeck</i>	226
--	-----

---

Session F1.4 Room 22A.110, 09:00–10:40

**Computational Advances in Data Analysis**

Chair: H.-J. Lenz

Hierarchical Clustering with the Maxim Method using DNA Computing <i>M. Cardona, M.À. Colomer, A. Zaragoza, J. Fibla</i>	227
---	-----

Nonlinear Time Series Control Charts <i>A. Messaoud, C. Weihs, F. Hering</i>	228
---	-----

OI — Object Integration on the Entity Level <i>H.-J. Lenz, M. Neiling</i>	229
--	-----

Dynaware-Tool for Mining <i>D. Thilagavathy</i>	230
--	-----

---

Session F1.5 Room 22A.111, 09:00–10:40

**Visualization**

Chair: P.J.F. Groenen

Multidimensional Scaling Algorithms for Large Data Sets <i>M.W. Trosset, P.J.F. Groenen</i>	231
--	-----

Interactive MDS-user Interfaces for Exploring Multidimensional Data Sets <i>P.J. Stappers</i>	232
---	-----

Visual Approaches for Classification in Large Databases <i>A. Wilhelm</i>	233
--	-----

Majorization for Symbolic Multidimensional Scaling <i>P.J.F. Groenen, O. Rodriguez, E. Diday, S. Winsberg</i>	234
--	-----

---

Session F1.6 Room 22A.112, 09:00–10:40

**Economics 2**

Chair: O. Opitz

Job Satisfaction as an Assessment Criterion of Labor Market Policy Efficiency — Classification Approach for International Comparisons <i>D. Kwiatkowska-Ciotucha, U. Zaluska</i>	235
--	-----

Classification-relevant Importance Measures for the West German Business Cycle <i>D. Enache, U. Garczarek, C. Weihs</i>	236
Classification of Open Pension Fund Functioning on Polish Market <i>P. Kowalczyk-Lizak, E. Poprawska</i>	237
Zur Modellierung und Prognose bei netzeffektbasierten Gütern <i>K.-H. Lüke, K. Ambrosi</i>	238

---

Session F1.7 Room 22A.120, 09:00–10:40

**Music Analysis 1**

Chair: C. Weihs

MusicMiner: A Databionic System to Organize Large Collections of Music <i>F. Mörchen, A. Ultsch</i>	239
Analysing the Classical Music Audience: Separating the Aging/Life Course Effect from the Cohort Effect <i>T.K. Hamann</i>	240
Modeling Melody Memory <i>D. Müllensiefen, C. Hennig</i>	241

---

Session F2.1 Room 22A.013, 11:00–12:15

**Data Mining and Explorative Multivariate Data Analysis 1**

Chair: L. D'Ambra / P. Giudici

Extended Centroid Method <i>V. Choulakian, L. D'Ambra, B. Simonetti</i>	242
Feature Selection in Predictive Genomic Mining <i>S. Figinì, P. Giudici</i>	243
Cultural Aspects and Social Stratification of Women: Overlapping Clustering and Multidimensional Scaling <i>M. Nakai</i>	244

---

Session F2.2 Room 22A.020, 11:00–12:15

**Text Mining 8: Classification 2**

Chair: G. Paaß

Improving Text Classification by Using Unlabeled Data <i>B. Novak, D. Mladenić, M. Grobelnik</i>	245
Evaluating the Performance of Text Classifiers on Real-World Press Archives <i>G. Paaß, H. de Vries</i>	246

Event Detection in Environmental Scanning:	247
News from a Hospitality Industry Newsletter	
<i>R. Wagner, J. Ontrup, S. Scholz</i>	

---

Session F2.3 Room 22A.105, 11:00–12:15

**Ranking, Multi-Label Classification, Preferences 1**

Chair: J. Fürnkranz / E. Hüllermeier

From Ranking to Classification: A Statistical View	248
--	-----

*S. Clemencon, G. Lugosi, N. Vayatis*

Monotone Function Decomposition	249
---------------------------------	-----

*V. Popova, J.C. Bioch*

On the Comparability and Construction of Reliability Measures for Dichotomous Ratings — A Unified Algebraic Approach	250
---	-----

*T. Ostermann, R. Schuster*

---

Session F2.4 Room 22A.110, 11:00–12:15

**Dissimilarities and Clustering Structures 1**

Chair: B. Fichet

Binary Clustering and Dissimilarities	251
---------------------------------------	-----

*F. Brucker*

Rescaling Proximity Matrix Using Entropy	252
--	-----

*S. Yokoyama, A. Okada*

One-Mode Two-Way Asymmetric Multidimensional Scaling by Using Joint Space Model	253
--	-----

*A. Okada, T. Imaizumi*

---

Session F2.5 Room 22A.111, 11:00–12:15

**PLS Path Modeling, PLS Regression and Classification 1**

Chair: N.C. Lauro / V.E. Vinzi

Model Building Strategies in PLS	254
----------------------------------	-----

*P. Hackl*

Classification in PLS Path Models and Local Model Optimisation	255
--	-----

*S. Squillacciotti*

PLS and GPA for Multiblock Analysis	256
-------------------------------------	-----

*M. Tenenhaus*

---

---

Session F2.6	Room 22A.112, 11:00–12:15
<b>Economics 3</b>	
Chair: O. Opitz	
Using Multivariate Statistical Methods to Identify Municipality Clusters	257
<i>D. Trocka</i>	
Spiralling in BTA Deep-hole Drilling — Models of Varying Frequencies	258
<i>N. Raabe, O. Webber, W. Theis, C. Weihs</i>	
Structural Differences in Rural Regions: An Analysis of Potential of Intermediate Central Communities	259
<i>A. Kubis</i>	

---

Session F2.7	Room 22A.120, 11:00–12:15
<b>Music Analysis 2</b>	
Chair: F. Mörchen	
Vibrato and Micromodulations as Cues for Timbre Segregation	260
<i>C. Reuter</i>	
Variable Steps and Stabilizing Intervals in our Tone Scale	261
<i>J.P. Fricke</i>	
Statistical Search for Melodic Prototypes	262
<i>K. Frieler, D. Müllensiefen, F. Riedemann</i>	

---

Session F3.1	Room 22A.013, 13:30–14:45
<b>Data Mining and Explorative Multivariate Data Analysis 2</b>	
Chair: L. D’Ambra / P. Giudici	
Dimensionality Reduction and Battery Applications	263
<i>I. Snihir, W. Rey, E. Verbitskiy</i>	
Using Subgroup Mining for the Refinement of Knowledge Systems	264
<i>M. Atzmueller, J. Baumeister, A. Hemsing, E.-J. Richter, F. Puppe</i>	
Restricted Co-Inertia Analysis	265
<i>P. Amenta, E. Ciavolino</i>	

---

Session F3.2	Room 22A.020, 13:30–14:45
<b>Text Mining 9: Retrieval Systems</b>	
Chair: W. Gaul	
Building Adaptive Knowledge Spaces by Combining Machine Learning Algorithms with Ontologies and Text Mining Technologies	266
<i>A.-C. Ngonga Ngomo, K. Böhm</i>	

Contexter — 267  
 A System for Visualisation of Large Collections of News Stories  
*M. Grobelnik, D. Mladenić*

Hot and Popular Terms in Search Queries 268  
*N. Schmidt-Mänz, W. Gaul*

---

Session F3.3 Room 22A.105, 13:30–14:45  
**Ranking, Multi-Label Classification, Preferences 2**  
 Chair: J. Fürnkranz / E. Hüllermeier

Ranking of Fault Models in Continuous Dynamic Systems 269  
 with Boosting  
*J.H. Rodriguez, C.J. Alonso*

On Active Learning in Multi-Label Classification 270  
*K. Brinker*

Ranking Syntactic Structures with Maximum Entropy Models 271  
*A. Eisele, C. Bering*

---

Session F3.4 Room 22A.110, 13:30–14:45  
**Dissimilarities and Clustering Structures 2**  
 Chair: B. Brucker

The Product of Metric Spaces 272  
 and the Product of Some Classification Structures  
*B. Fichet*

A Multidimensional Scaling Method 273  
 as a Method of Information Compression  
*T. Imaizumi*

Active Learning Using Expected Likelihood-based Sampling 274  
 for Classifiers based on Mixture Models  
*B.H. Iswanto, K. Obermayer*

---

Session F3.5 Room 22A.111, 13:30–14:45  
**PLS Path Modeling, PLS Regression and Classification 2**  
 Chair: N.C. Lauro / V.E. Vinzi

PLS Path Modeling: Optimality Properties of Latent Variables 275  
*M. Hanafi*

Assessing Unidimensionality of a Set of Variables — 276  
 Application to PLS Path Modeling  
*K. Sahmer, M. Hanafi, E.M. Qannari*

Multivariate Partial Robust M Regression 277  
*S. Serneels, C. Croux, P. Filzmoser, P.J. Van Espen*

---

Session F3.6 Room 22A.112, 13:30–14:45

**Economics 4**

Chair: O. Opitz

Analysis of the Economic Development of Districts in Poland  
 as a Basis for the Framing of Regional Policies 278

*M. Rozkrut, D. Rozkrut*

Clustering of Economic Time Series 279

*D. Rozkrut*

The Classification of Local and Branch Labour Markets  
 in the Upper-Silesia 280

*W. Hantke*

---

Session F3.7 Room 22A.120, 13:30–14:45

**Music Analysis 3**

Chair: D. Müllensiefen

klaR - Classification Tools for R: 281

Case Study in a Music Sight Reading Task

*U. Ligges, K. Luebke, N. Raabe, R. Kopiez, C. Weihs*

Connectionist Approaches to Music Information Retrieval 282

*S. Vembu*

Parameter Optimization in Automatic Transcription of Music 283

*C. Weihs, U. Ligges*

---

Plenary F Room 16.H5, 14:50–15:35

**SolEuNet: Selected Data Mining Techniques and Applications**

Speaker: N. Lavrač

Chair: W. Gaul 284

---

## Bibliothekarisches Programm: Inhaltserschließung 2005

**Mittwoch, 9. März 2005**

**Arbeitsgruppe Bibliotheken — bibliothekarische Fortbildung**

Raum: Universitätsbibliothek

10:00	Begrüßung, Einführung	Dr. Hans-Joachim Hermes, Chemnitz
10:15	Leitreferat 1: Informationswesen / Sacherschließung in Deutschland. Lösungsansätze für eine Welt der polyzentrischen Informationsversorgung	Prof. Dr. Jürgen Krause, Bonn
11:15	Leitreferat 2: Informationslandschaft Europa (European Information Subject Access Panorama)	Patrice Landry, Bern
12:15	Suchverhalten im Web: Empirische Ergebnisse	Dipl.Wi.-Ing. Nadine S. Schmidt-Mänz, Karlsruhe
14:00	In eigener Sache: Wahl des Vorsitzenden der Arbeitsgruppe Bibliotheken. Festlegungen zur Struktur der AG BIB	AG BIB
14:15	Neu in der ETH-Bibliothek: Computerlinguistik im Einsatz	Klaus Loth, Zürich
15:00	Neues zur Sacherschließung im Südwestverbund	Stefan Wolf, Konstanz
15:45	Neues zu Portalen	Kerstin Zimmermann, Wien
16:30	Abschlussdiskussion	Moderation: Dr. Hans-J. Hermes
anschl.	Rundgang durch die Universitätsbibliothek	BDirektor Dr. Ekkehard Oehmig

## Bibliothekarisches Programm: Inhaltserschließung 2005

**Donnerstag, 10. März 2005**

**Bibliothekarische Sektion**

Raum: Universitätsbibliothek

10:00	Arbeitsgruppe Dezimalklassifikationen Öffentliche Sitzung:  Anwenderberichte: DK in Freiberg, Leipzig und Mittweida  Literaturbericht zu den Dezimalklassifikationen	Moderation: Dr. Bernd Lorenz  Dagmar Rothe, Leipzig; Dr. Andreas Stumm, Freiberg; N.N., Mittweida  Dr. Bernd Lorenz, München
11:15	DDC Deutsch. Vom Projekt zur Dienstleistung	Magda Heiner-Freiling, Die Deutsche Bibliothek, Frankfurt
12:00	Klassifikationen, Cross-Konkordanzen, virtuelle Fachbibliothek "GEO-LEO"	Dr. Andreas Stumm, Freiberg und Dr. Norbert Pfurr, Göttingen
14:30	Krünitz-Projekt: Erschließung der Enzyklopädie (1773-1858, 242 Bde.) mit DDC	Dr. Hans-Ulrich Seifert, Trier
15:15	Das Informationsportal der UB Magdeburg	Eckhard Blume und Gebhard Engelmann, Magdeburg
anschl.	Rundgang durch die Universitätsbibliothek	BDirektor Dr. Ekkehard Oehmig