



# International Federation of Classification Societies Newsletter

**President:** David Hand

Number 32

November 2006

**Editor:** André Hardy

**The International Federation of Classification Societies**, founded in 1985, is composed of:

Associação Portuguesa de Classificação e Análise de Dados, British Classification Society, Central American and Caribbean Society of Classification and Data Analysis, Classification Society of North America, Gesellschaft für Klassifikation, Irish Pattern Recognition and Classification Society, Japanese Classification Society, Korean Classification Society, Société Francophone de Classification, Società Italiana di Statistica, Statistical Society of Slovenia, Vereniging voor Ordinate en Classificatie and Section of Classification and Data Analysis of the Polish Statistical Society.

The IFCS is a non-profit, non-political scientific organization, the aims of which are to further classification research. Among other activities, the IFCS organises a biennial conference, and supports the "Journal of Classification" and the journal "Advances in Data Analysis and Classification (ADAC)".

## **IFCS Homepage:**

<http://www.classification-society.org/>

Designed and maintained by David Dubin.

The website contains among others: the IFCS Constitution and By-Laws, the IFCS newsletters, and pointers to the websites of the member societies.

## **Contents**

Words from the IFCS President  
News from the IFCS Scientific Secretary  
A new Member Society of the IFCS  
News from the VOC  
News from the GfKI  
News from the JCS  
News from SKAD  
News from CSNA  
Short news from CLAD  
News from the Statistical Section of the Korea Statistical Society  
News from the SFC  
IFCS Transversal Group on Symbolic Data Analysis  
EGC 2007  
European Workshop on Datastream Analysis  
Fourth International Meeting on Statistical Implicative Analysis (ASI 4)  
CARME 2007  
IASC International Conference on Statistics for Data Mining, Learning and Knowledge Extraction  
Compstat 2008  
4<sup>th</sup> World Conference on Computational Statistics & Data Analysis of IASC and 6<sup>th</sup> Conference of Asian Regional Section of IASC  
IASC-ERS Summer School on Data Mining and Statistical Learning Techniques  
New Journal on data analysis and classification  
Short book review

## **Words from the IFCS President November 2006**

As is reported elsewhere, the 10th IFCS Conference took place in Ljubljana in July. The organisers, the Statistical Society of Slovenia and the Faculty of Social Sciences, University of Ljubljana, did a superb job. The innovation of

arranging that the conference badge doubled as a bus pass, enabling free bus travel, was brilliant, and made at least my journeys to the conference venue effortless!

The scientific programme covered a huge range of topics. The extraordinary breadth of classification science is illustrated by the fact that applications addressed by the various speakers included similarity of melodies, analysis of financial data, fraud, marketing and advertising, patent applications, internet data, drug usage, food science, cancer screening, social networks and relationships, bibliometrics, symbolic data, authorship attribution, customer value management, and safety. Interwoven with these were a large number of theoretical and methodological advances. This range demonstrates how important classification is, as well as how dynamic the modern discipline is. Each new application comes with its own particular challenges, and the solutions, once found, then readily find application in other areas. Large  $d$ , small  $n$  problems are a perfect illustration: much of the work on such problems has been motivated by applications in the biological microarray domain, but the tools are also important elsewhere. Such problems have become an important special focus of methodological research.

The conference organisers kindly allowed me to organise a 'President's invited session', in which I invited three speakers who, I felt, had something to say which would be of general interest to the conference delegates. I believe my three speakers, Fionn Murtagh, Henk Kiers, and Friedrich Leisch fulfilled my expectations. Certainly they stimulated some interesting debates. Adrian Raftery, the keynote plenary speaker, addressed model-based clustering, in particular in the context of social networks, an inspired choice since Ljubljana is known for its research in this area.

In addition to the above, the conference featured a rich series of presentations, typically with three parallel streams and with many invited speakers - to the extent that there often a problem of choice arose.

As has been reported by the Scientific Secretary, for administrative reasons, the next conference will be held in 2009, in Dresden, rather than in

2008. This will put us back on the track disrupted by the Kobe earthquake, holding the conference in even numbered years. I look forward to receiving bids to hold the 2011 meeting.

The report of the Scientific Secretary lists other classification societies currently in discussions on joining the IFCS. Broadening our base will help to bring the international classification community together.

In this vein, I welcome suggestions for other ways in which the IFCS can continue to serve our scientific community more effectively. As a minor innovation in this regard, we have decided to introduce the innovation of having brief reviews of books relevant to classification researchers in the Newsletter. We hope you find them interesting and valuable.

David J. Hand  
IFCS President

## **News from the IFCS Scientific Secretary**

Dear Colleagues and Friends of the Classification,

The 10<sup>th</sup> Conference of the IFCS took place in Ljubljana (Slovenia) last 25-29 July. IFCS'06 was organized by the Statistical Society of Slovenia and the Faculty of Social Sciences, University of Ljubljana. The Conference was attended by 156 participants. More information on the scientific programme as well as some pictures from the scientific sessions and the social events is still available on the site: <http://vlado.fmf.uni-lj.si/info/ifcs06/>.

The IFCS Executive Committee and Council meetings took place during IFCS 2006 in Ljubljana. Several issues and projects related to the scientific and administrative life of IFCS were discussed during these meetings. The Minutes of the Council meeting have been sent to the representatives of all IFCS member societies. Taking into account also later discussions, here are the main news from the Executive Committee and Council:

## **IFCS 2009 Conference**

The next IFCS meeting will be organized by GfKl and will take place on March 13-18, 2009 in Dresden, Germany.

## **IFCS 2011: Future Candidatures**

Taking into account that the next IFCS conference will take place in 2009, bids are invited for the IFCS conference to be held in 2011. Potential ideas shall be forwarded to David Hand (IFCS President) for further discussion.

## **New IFCS Member Societies**

As already announced in the previous Newsletter (number 31 of May 2006), a very important project which IFCS is currently engaged with is represented by the inclusion of new member societies in the Federation. The contacts with the Greek Society of Data Analysis and the Chinese Association for Applied Statistics are still ongoing and will most likely lead to a formal submission of their IFCS membership applications. Meanwhile, the Statistical Society of Slovenia as a whole including SLAIS (Slovenian Artificial Intelligence Society) submitted a formal IFCS membership application. The IFCS Council will have to express a formal vote by the end of the year.

## **IFCS co-sponsored conferences**

The CARME (Correspondence Analysis and Related MEthods) Transversal Group of IFCS will hold its conference from June 25 to 27, 2007 in Rotterdam (<http://www.carme-n.org/carme2007/>).

The KNEMO'06 (KNowledge Extraction and MOdelling) International Workshop took place in Anacapri (Italy), from September 4th-6th ([www.knemo.unina.it](http://www.knemo.unina.it)). The Workshop, a satellite meeting of the COMPSTAT'06 Conference, was organized by the Department of Mathematics and Statistics of the University of Naples "Federico II" with the scientific sponsorship of the IASC (International Association for Statistical Computing), the INTERFACE Foundation of

North America and the IFCS. 104 participants attended the Workshop.

Vincenzo Esposito Vinzi  
IFCS Scientific Secretary

## **A new Member Society of the IFCS**

After a formal vote by the IFCS Council, I am happy to inform that the Statistical Society of Slovenia (SSS) has been accepted as an official member society of the IFCS.

Vincenzo Esposito Vinzi  
IFCS Scientific Secretary

## **News from the VOC**

The autumn meeting of the VOC on November 3, 2006 at Leiden University was a successful event with about 50 attendants. The theme 'View on Changes: models for longitudinal data' was discussed from different perspectives. Geert Verbeke (Catholic University of Leuven) presented an illuminating example of mixed models for prediction in a medical application using multivariate longitudinal profiles. Reinoud Stoel (University of Amsterdam) discussed the consequences of using inequality constraints for the likelihood ratio test in structural equation modeling. Ellen Hamaker (Utrecht University) talked on time series analysis to study psychological processes. Cees Elzinga (Free University) treated various metric representations of categorical time series. Mark de Rooij (Leiden University) discussed his gravitational models for the analysis of change. Paul van Geert (Groningen University) closed our meeting with a lively overview of the use of dynamic system models for development. The up-to-date information on the activities of the VOC is available via [www.voc.ac](http://www.voc.ac).

Marieke Timmerman

## News from the GfKI

### 31<sup>st</sup> GfKI Annual Conference

The German Classification Society GfKI will hold its 31<sup>st</sup> Annual Conference on **Data Analysis, Machine Learning, and Applications** at the University of Freiburg (Germany) from March 7 – 9, 2007.

The scientific program will include invited talks by

- Ricardo Beaza-Yates (Universitat Pompeu Fabra, Spain)
- Shai Ben-David (University of Waterloo, Canada)
- Paul Eilers (Leiden University Medical Centre, Netherlands)
- Anuska Ferligoj (University of Ljubljana, Slovenia)
- Eugeniusz Gatnar (Karol Adamiecki University of Economics, Poland)
- Christian Hennig (University College London, London, UK)
- Vincent Lemaire (France telecom R&D, France)
- Tobias Scheffer (Humboldt University, Berlin, Germany)
- Donatella Vicari (University of Rome “La Sapienza”, Italy) and
- Simon Wood (University of Bath, UK)

as well as sessions with invited and contributed talks from a broad range of topics.

The conference addresses current research on **Data Analysis, Machine Learning, Data Mining, Pattern Recognition**, and **Applied Statistics** as well as **applications** to a broad range of areas with special emphasis on interdisciplinarity and the interaction between theory and practice. Topics include, but are not limited to:

#### Theory and Methods:

- Supervised Classification, Discrimination and Pattern Recognition
- Cluster Analysis and Similarity Structures
- Classification and Regression
- Frequent Pattern Mining
- Data Visualization and Scaling Methods
- Exploratory Data Analysis and Data Mining

- Mixture Analysis in Clustering
- Computational Intelligence
- Knowledge Representation and Knowledge Discovery
- Statistical Relational Learning
- Online Algorithms and Data Streams
- Analysis of Time Series, Longitudinal and Panel Data
- Tools for Intelligent Data Analysis
- Data Preprocessing and Information Extraction
- Typing for Modeling

#### Applications:

- Marketing and Management Science
- Banking and Finance
- Business Intelligence and Personalization
- Data Analysis in Retailing
- Econometrics and Operations Research
- Image and Signal Analysis
- Biostatistics and Bioinformatics
- Medical and Health Sciences
- Text Mining, Web Mining, and the Semantic Web
- Statistical Natural Language Processing
- Linguistics
- Subject Indexing and Library Science
- Statistical Musicology
- Archaeology and Archaeometry
- Psychology
- Data Analysis in Higher Education

The program will be complemented by contributed sessions from classification societies of other countries, e.g.

- Latent Class Models for Classification (SIS, Italy),
- Classification and Models for Interval-valued Data (SIS, Italy),
- Selected Problems in Classification (SKAD, Poland),
- Recent Developments in Multidimensional Data Analysis - Between Research and Practice (SIS, Italy).

Furthermore, there will be a doctoral workshop for PhD students working on the conference topics and special "Junior Sessions" (with their own best paper award) on data analysis and classification for young researchers.

Participants intending to present a talk are invited to submit an abstract via the conference website **until December 1, 2006**. All abstracts will be reviewed and the accepted ones will be distributed to the conference participants. A post-conference proceedings will be published in the Springer Series "**Studies in Classification, Data Analysis, and Knowledge Organization**". The best paper will be awarded the **GfKI best paper award**.

Conference languages are English and German.

#### **Important dates:**

- **Dec. 1, 2006:** Deadline for abstract submission (extended by request)
- Jan. 12, 2007: Notification of acceptance for talks
- Mar. 6, 2007: GfKI Doctoral Workshop
- **Mar. 7-9, 2007: 31st GfKI Conference**
- Mar. 23, 2007: Deadline for full paper submission (for the post-conference proceedings)
- May 31, 2007: Notification of acceptance for papers
- June 22, 2007: Deadline for camera-ready papers.

Detailed information about the conference can be found on the conference website

<http://www.informatik.uni-freiburg.de/gfki/>

If you have any further questions concerning the conference, please do not hesitate to contact the local organizing team:

[gfki@informatik.uni-freiburg.de](mailto:gfki@informatik.uni-freiburg.de)

Lars Schmidt-Thieme,  
Hans Burkhardt,  
Reinhold Decker

## **28th Fall Meeting of the working group AG-DANK.**

The 28th Fall Meeting of the Working Group for Data Analysis and Numerical Classification (AG-DANK) of the GfKI took place at the Department of Statistics of the University of Dortmund, invited by Professor Claus Weihs. The interdisciplinary meeting was attended by more than 30 participants, among them five from abroad. The Friday session and part of the Saturday session were devoted to this year's focal topic, "ensemble methods." We had eight talks on this topic, ranging from the combination of classifiers, bagged logic regression, and random forests to kernel combination. The invited speaker, Professor Peter Bühlmann of the Swiss Technical University in Zürich (ETH), gave a presentation on "Boosting and twin boosting for high-dimensional data." The remaining time on Saturday was devoted to talks from other domains of data analysis and classification. Lists of references of most of the talks are posted in the internet

<http://stoch.fmi.uni-passau.de/agdank/dortmund2006/index.php3>

Gunter Ritter

## **News from the JCS**

### **The 23rd Annual Research Meeting of the Japanese Classification Society**

The 23rd Annual Research Meeting of the Japanese Classification Society (JCS) will be held at the RENAISSANCE CENTER of Tama University on November 25-26, 2006.

The official Web site is  
<http://www.ir.rikkyo.ac.jp/~jcs2006/> .

Japan and Korea Joint invited session is included in the meeting.

JCS invites several members from the Korean Classification Society.

## **2. The 22nd Annual Research Meeting of the Japanese Classification Society**

The 22nd Annual Research Meeting of the Japanese Classification Society was held at the Tokyo Rikkyo (St. Paul's) University on December 6, 2005. The speakers and the titles of their talks are as follows:

### **Visualization Methods to Explore Clustering Results**

Yoshiro Yamamoto \*, Hiroshi Yadohisa \*\*  
\* Tokai University, \*\* Doshisha University

### **Clustering Considering the Relation of Fuzzy Classification Structures**

Tomoyuki Kuwata, Toshiya Shijo, Mika Sato-Ilic  
University of Tsukuba

### **Classification of the Cultural Patterns using Property Values**

Yasunari Kono  
Rikkyo (St. Paul's) University

### **Analysis of Change Price in Stocks and Foreknowledge an Earthquake**

Akiyoshi Kondou  
System Control Co., Ltd.

### **Knowledge Extraction from Word-Of-Mouth Data in Web**

Huang Chih Chieh, Takao Ishizuka  
Asia University

### **Statistical Model for Estimating Assymmetric Price Response of each "ONE" Customer: A Case Study Developing Marketing Information Systems for Retail Company**

Kotaro Shibuya  
Saison Information Systems Co., Ltd.

### **Difference in Influence of Consumptive Awareness that Affects the Sense of Spending by Sex, using Time-series Data: An Attempt using Pre-survey Data with Multiple Indicators Model**

Nobuyuki Otake  
Marketing Service Co., Ltd.

### **Secondary Analysis and Security**

Kiyoshi Matsui  
TOKYO CRO Inc.

### **Analyzing a Proximity Matrix which Represents Uncertainty**

Satoru Yokoyama \*, Akinori Okada \*\*  
\* Keio University, \*\* Rikkyo (St. Paul's) University

### **MDPREF Analysis of Chinese Characters Used in Given Names from 1912 through 2004**

Mayuko Etou, Akinori Okada  
Rikkyo (St. Paul's) University

### **Analyzing Asymmetric Preference Relationships among Sports by Utilizing Asymmetric Cluster Analysis**

Noriko Sakai, Akinori Okada  
Rikkyo (St. Paul's) University

### **One-mode Three-way Data Analysis using Variable Weights Model**

Atsuh Nakayama  
Rikkyo (St. Paul's) University

### **An Extension of the CHAID Tree-based Segmentation Algorithm to Multiple Dependent Variables**

Jay Magidson  
Statistical Innovations Inc

### **Latent Structural Modeling for Multi-categorical Segments Mapping: Application for Real Data**

Satoko Kojima\*, Gen Tamura\*\*, Kazuo Shigemasu\*

\*The University of Tokyo, \*\*Video Research Ltd.

### **The Latent Class Model for Simultaneous Two Dimensional Analysis**

Naoko Sakurai \*, Kazunori Yamaguchi \*\*  
\* Tokyo University of Information Sciences, \*\* Rikkyo (St. Paul's) University

### **Item Parceling and Model Fit in Confirmatory Factor Analysis**

Kensuke Okada, Takahiro Hoshino, Kazuo Shigemasu  
The University of Tokyo

Yoshiro Yamamoto

## News from SKAD

The Fifteenth Annual Conference of Section of Classification and Data Analysis of Polish Statistical Association took place on September 20<sup>th</sup> to 22<sup>nd</sup> in Wisła. The conference was organized by the Department of Statistics of Katowice University of Economics. Almost 100 persons participated in the conference, presenting 73 papers and 10 posters.

Keynote speaker was Professor David J. Hand, who presented talk: "Size matters: measurement and science".

During the conference the elections took place. The Council of SKAD was elected for years 2007-2008:

Krzysztof Jajuga – chairman

Andrzej Sokołowski – vice-chairman

Eugeniusz Gatnar – secretary

Zdzisław Hellwig – member

Kazimierz Zajac – member

Waldemar Tarczyński – member

Marek Walesiak – member

Joanicjusz Nazarko – member.

It was decided that next annual conference will take place in September 2007.

Krzysztof Jajuga

## News from CSNA

### Call for Participation

#### The 2007 Meeting of the Classification Society of North America

##### Finding Structure in Complex Data

Extracting useful information from large and complex machine-readable data sets is a problem faced by people in nearly every area of commerce, manufacturing, government, and in every academic discipline and science. Creating an environment conducive to solving such problems is not a mystery, but it is difficult since it requires

assembling a community with broad collective expertise in mathematics, statistics, computing, and a wide variety of different application areas.

For the past thirty years the Classification Society of North America (CSNA) has gathered just such a community at its annual meetings. CSNA members contribute a wide range of backgrounds in mathematics, statistics, computing, information and decision sciences, in addition to expertise in business, the physical, social, life, and earth sciences. We are brought together by a common interest in the shape of data analysis, extraction and summarization problems, and with the goal of making sense of those commonalities despite differences in terminology, notation, and the disparate professional literatures in which we read and publish. Our meetings offer the participant a kind of productivity that can't be matched at the discipline-specific, tool-specific, and problem-specific meetings that we attend at other times of the year.

### CSNA 2007

The 2007 CSNA meeting will take place June 7-10, 2007, at the University of Illinois at Urbana-Champaign in Urbana, Illinois. It is organized by David Dubin and Carolyn Anderson. The meeting is sponsored by the National Center for Supercomputing Applications and the UIUC Graduate School of Library and Information Science. We invite proposals for contributed paper sessions in all areas relating to classification, clustering and data analysis.

The types of problems that interest us are often described as machine learning, pattern recognition, or data mining problems, or with reference to the application area such as information retrieval, molecular bioinformatics, authorship attribution, market segmentation, psychometrics, social networks, and so on. One can get a better sense of our diverse application areas and computational techniques by browsing the abstracts from previous meetings (2006, 2005, 2004, 2003).

CSNA's keynote speaker will be Michael J. Kurtz of the Smithsonian Astrophysical Observatory. Dr. Kurtz will be speaking on The Astronomical Information Network.

CSNA meeting sessions are traditionally informal, with presentations aimed at soliciting critique, discussion, and suggestions for improving the work. In keeping with this

atmosphere, we distribute abstracts of presented papers, but do not publish a proceedings.

### **CSNA and Digital Humanities**

This year CSNA follows directly after the 2007 Digital Humanities meeting, which will take place at the same venue. Although CSNA sessions will reflect the usual range of disciplines, our meetings will intersect on June 7 with a Joint Workshop on Data Analysis and Research in the Humanities. In addition, DH 2007 and CSNA 2007 will have a mutual registration agreement, by which participants in either meeting will be able to attend sessions at both. So plan to come to Urbana early for some or all of the Digital Humanities meeting starting June 2nd. Further information on DH 2007 can be found at <http://www.digitalhumanities.org/dh2007/>

### **Student Poster Competition**

Students enrolled full time in a graduate or undergraduate program are eligible to participate in our student research poster competition. First prize (to be announced at the meeting) is a \$500 travel stipend. Posters will be judged by a panel of data analysis experts on research originality, technical merit, and effectiveness of presentation. There is a limit of one entry per student, and one author per entry. Poster topics can be in any area of classification research.

Students wishing to enter the competition should submit a 1-2 page extended abstract of the work by email to the proposal address below no later than March 2, 2007. Selected entries will be invited to exhibit posters for the judging panel at the meeting.

### **Submitting your proposal**

Proposals for CSNA papers are due March 2, 2007 and should include a title, author affiliation and contact information, and an abstract of no more than 200 words. This information should be sent as either a plain text or LaTeX attachment by email to [csna2007@classification-society.org](mailto:csna2007@classification-society.org). In the event of difficulties using the address, contact David Dubin at the University of Illinois either by email ([ddubin@uiuc.edu](mailto:ddubin@uiuc.edu)) or by telephone (217-244-3275).

Proposals for work at the intersection of data analysis and humanistic scholarship should indicate whether they wish to be considered for case study discussion at the Joint DH/CSNA workshop. Student poster competition proposals are limited to two pages.

Join us in Urbana

Registration, travel, and schedule information for CSNA 2007 will be posted as it becomes available at <http://www.classification-society.org/csna/csna07.html> Further information about the society can be found at <http://www.classification-society.org/csna/> CSNA is a member of the International Federation of Classification Societies (IFCS).

David Dubin  
[ddubin@UUIC.EDU](mailto:ddubin@UUIC.EDU)

Stanley L. Sclove  
Secretary of CSNA  
[sclslove@uic.edu](mailto:sclslove@uic.edu)

### **BENEFITS OF CSNA MEMBERSHIP**

For the year 2007, members of IFCS societies can join CSNA at the rate of \$60 US.

Members receive the Journal of Classification. Also, CSNA's bibliographic "Service" CD (Classification Literature Automated Search Service), edited by Fionn Murtagh, is included with all memberships. This CD includes clustering and classification bibliographies with over ten thousand references from 1994 to 2005, and several out-of-print classic books on classification / clustering in their entirety, including those of Bock, Dubes and Jain, Hartigan, Lerman, Murtagh, and van Rijsbergen. The CD is mailed with the first issue of the Journal each year.

The second issue of the Journal went to press toward the end of October and should be received by subscribers in November.

Springer, the publisher of the Journal of Classification, offers their books to CSNA members at a 20% discount. Upon paying CSNA dues, persons will receive the requisite pass codes from the CSNA Secretary.

CSNA members can access back issues of Journal of Classification on line via Springer's Website.

Stanley L. Sclove  
Secretary of CSNA  
[sclslove@uic.edu](mailto:sclslove@uic.edu)

## Short news from the “Associação Portuguesa de Classificação e Análise de Dados”, CLAD

- **CLAD** is co-sponsoring the **ISI 2007** meeting of the International Statistical Institute, that will be held in Lisbon (<http://www.isi2007.com.pt/>) and also sponsors the Invited Paper Meeting IPM 90 on Analysis of Complex Data in Health Sciences.
- **CLAD** and **LEAD** (<http://lead.paginas.sapo.pt/>) are co-sponsoring the two following **ISI 2007** satellite meetings:
  - **ISBIS 2007** - International Symposium on Business and Industrial Statistics, that will be held in the “Universidade dos Açores” (<http://www.isbis2007.uac.pt/intro.php>)
  - **ICSMRA 2007** - International Conference on Statistical Methods for Risk Analysis, will be held in the Universidade de Lisboa (<http://www.fm.ul.pt/ICSMRA2007>).
- The **JOCLAD2007** annual meeting will take place at the Instituto Superior de Engenharia do Porto, on the 8<sup>th</sup>-10<sup>th</sup> February 2007.  
([www.JOCLAD2007.isep.ipp.pt](http://www.JOCLAD2007.isep.ipp.pt))

Helena Bacelar-Nicolau

## News from the Statistical Classification Section of the Korea Statistical Society.

New staff of the Korea Statistical Classification Section of the Korea Statistical Society.

- Chief

Sung-Soo Kim, Professor, Department of Information Statistics, Korea National Open University.  
e-mail : [sskim@knou.ac.kr](mailto:sskim@knou.ac.kr)

- Secretary

Hyunjoong Kim, Associate professor, Department of Applied Statistics at Yonsei University.  
e-mail : [hkim@yonsei.ac.kr](mailto:hkim@yonsei.ac.kr)

- Period : 2006. September 01 - 2008. August 31.

Sung-Soo KIM  
Chief of the Classification Section of the Korea Statistical Society.

## News from the SFC

### Election

François-Xavier Jollois (Paris 5 University, France) has been elected as Editor of the SFC Bulletin.

André Hardy (University of Namur, Belgium) has been reelected as Secretary of the SFC.

### SFC 2006

The SFC'2006 meeting took place in the nice city of Metz (France), September 6-8, 2006.

The objective was to gather researchers and application developers from a wide range of classification related areas such as machine learning, pattern recognition, and data visualization. The program included communications on the latest developments in micro array data, text mining, pattern recognition,

visualization, data mining and optimization. The meeting was attended by more than 70 participants. The invited speakers were Andrea Cerioli (University of Parma, Italy), Joachim Buhmann (Institute of Computational Science, Zurich, Switzerland), Patrick Gallinari (Université Pierre et Marie Curie, Paris, France), Mark Girolami (University of Glasgow, Scotland), Iven Van Mechelen (University of Leuven, Belgium), Jean-Michel Poggi (Université Paris-Sud, France) and Adrian Raftery (University of Washington, USA). A selected number of accepted papers will be published in the French journal RNTI (<http://www.antsearch.univ-tours.fr/rnti/>) published by [Cépaduès Editions](#).

André Hardy,  
Secretary of SFC

## **SFC 2007**

The SFC 2007 Conference will be held in Paris (France) on September 5-7, 2007. It will be organized by Telecom Paris (Ecole Nationale Supérieure des Télécommunications de Paris) which is situated in the heart of Paris.

Among the standard topics of this conference, the focus in 2007 will be on

- Graphs and classification
- Bioinformatics
- Data stream clustering
- Social networks, web
- Textual analysis
- Symbolic data analysis.

Olivier Hudry will chair the program committee. I. Charon, G. Hébrail and O. Hudry will be in charge of the organization of the conference.

Conference web site:  
<http://www.enst.fr/SFC2007>

Georges Hébrail

## **IFCS Transversal Group on Symbolic Data Analysis**

The IFCS Transversal Group on Symbolic Data Analysis continues to develop its activities to promote and disseminate recent research on this topic. In this context, a Tutorial on “Knowledge Mining by Symbolic Data Analysis and the Sodas Software” has been organized during the last IFCS Conference, in Ljubljana, which allowed to present recent developments of the analysis of symbolic data. The conference also comprehended an invited session on “Symbolic Data Analysis”, organized by R. Verde and Y. Lechevallier. Members of the transversal group also promote and organize sessions within the framework of other international conferences: an invited session with title “From the Statistic of Standard or Complex Data to the Statistic of Knowledge: Symbolic Data Analysis” was organized by E. Diday during the COMPSTAT-2006 conference in Rome, last August. Invited Paper Meetings as well as Special Topic Contributed Paper Meetings on related topics are foreseen for the next ISI Session to take place in Lisbon in Summer 2007.

Recently, an advanced course on “Knowledge extraction by Interval Data Analysis”, organized by *Nova Universitas*, took place in Caserta, Italy. The website of the IFCS Transversal Group on Symbolic Data Analysis is presently under construction and will be opened in the beginning of 2007.

Paula Brito

## **EGC 2007**

**Knowledge extraction and management  
(Extraction et Gestion des Connaissances  
2007)**

January 2007, 23-26, Namur, Belgium  
<http://www.info.fundp.ac.be/egc2007>

The 7th French Conference "Extraction et Gestion des Connaissances" (EGC'2007) will take place in the Institut d'Informatique, Facultés

Universitaires Notre-Dame de la Paix, Namur, Belgium, January 2007, 23-26. This conference is devoted to theories, methods and applications of Data Mining, Machine Learning, Knowledge Discovery in Data Bases, Knowledge Management and engineering. Its aim is to gather all researchers and company managers in this area in order to present work of high quality and exchange new ideas.

EGC'2007 includes tutorials, workshops, oral presentations, posters, and software demonstrations. The names of the invited speakers and the list of the accepted papers are already available on the website of the conference.

Two special scientific prizes of 1500 euros each will be respectively awarded to the best academic and industrial papers.

The proceedings of EGC'2007 will appear in a special issue of the French journal RNTI (Cépaduès publisher).

More information (registration, ...) on the website.

Monique Noirhomme and Gilles Venturini,  
(Local chair and program chair).

## **EUROPEAN WORKSHOP ON DATASTREAM ANALYSIS**

March, 14-16, 2007

Belvedere di S. Leucio, Caserta - Italy

<http://www.jsda.univa2.it/WDSA/home.html>

### **CALL FOR PAPERS**

This European Workshop seeks to bring together researchers working on data base, machine-learning, decision rules, decision trees, association rules, filtering, data mining, clustering, visualization techniques, etc. from data streams and related themes. While the primary objective is to provide an avenue for dissemination of research results and works in progress, we also seek to discuss the future development of knowledge discovery in data streams, and how the research will benefit applications.

### **Structure of the WORKSHOP**

The Workshop program includes:

- 2 tutorials by J. Gehrke (USA) and G. Cormode (USA)
- 3 invited talks by J. Aguilar (Spain), P. Giudici (Italy), O. Rana (United Kingdom)
- 3 invited sessions; 3 contributed sessions and a poster session.

Rosanna Verde.

### **Fourth international Meeting on Statistical Implicative Analysis (ASI 4) University of Castellón, Spain October 18-21, 2007**

Statistical Implicative Analysis (SIA) is a non symmetrical multivariate data analysis method involving several type of data: binary, categorical, numeric, interval, fuzzy, ... Data is structured and new knowledge is provided under the form of rules and generalized rules among variables. The origin of SIA is the stochastic modelization of the so called quasi-implication: "whenever the variable (or conjunction of variables)  $a$  is observed in the population, then, in general, so is the variable (or conjunction of variables)  $b$ ." The role played by specific data in the results of the analyses can be examined through the notions of contribution and typicality. All SIA results are easily displayed and interpreted through software CHIC (Classification Hiérarchique Implicative et Cohésitive).

Structured in 1-day tutorial talks and 2-day theoretical and applied communications, topics include ASI statistical models, duality variable-individual spaces, application on didactics, psychology, sociology, economy, biology, etc.

Communications to be sent:

[rencontreASI4@polytech.univ-nantes.fr](mailto:rencontreASI4@polytech.univ-nantes.fr)

Deadline is fixed to March 1, 2007

More information on <http://www.asi4.uji.es>.

Pres. Sci. Comm.: Régis Gras, LINA, Université de Nantes (France)

Régis Gras



## First Call for Papers

Correspondence Analysis and Related Methods,  
Erasmus University Rotterdam, The Netherlands  
Monday June 25 to Wednesday June 27, 2007  
<http://www.carme-n.org/carme2007>

The objective of this conference is to spotlight the very latest research in correspondence analysis, and discuss future developments. We aim to bring together theoretical and applied researchers in all the areas where correspondence analysis is currently used, notably sociology, psychology, economy, education, ecology, archeology, geology, linguistics, philosophy, genetics, biomedical research, health economics, marketing and management. Interdisciplinary contributions will be particularly welcome.

This conference is the fifth in a series of conferences every four years on correspondence analysis, multidimensional graphical techniques, and the analysis of large sets of categorical data:

- May 1991: Correspondence Analysis in the Social Sciences, Cologne (Germany);
- May 1995: Visualisation of Categorical Data, Cologne (Germany);
- May 1999: Large Scale Data Analysis, Cologne (Germany);
- July 2003: Correspondence Analysis and Related Methods, Barcelona (Spain).

Themes of the conference include all forms of correspondence analysis and related fields, including visualisations of categorical data:

- Simple correspondence analysis
- Multiple correspondence analysis
- Joint correspondence analysis
- Multiway correspondence analysis

- Canonical correspondence analysis
- Nonsymmetrical correspondence analysis
- Dual scaling
- Optimal scaling
- Homogeneity analysis
- Multidimensional scaling of categorical data
- Biplots of categorical data
- Visualisation of compositional data
- Correspondence analysis in the social sciences
- Correspondence analysis in ecology and the environmental sciences
- Correspondence analysis in the health sciences
- Correspondence analysis in marketing research and management

## Preliminary Program

- Monday, June 25: Conference.
- Tuesday, June 26: Conference.
- Wednesday, June 27: Conference and Conference dinner.

## Dates and fees:

- March 31, 2007: Submission abstract.
- Around April 15, 2007: Notification of acceptance of conference papers.
- Before April 30, 2007: Early registration fee of € 120.
- After April 30, 2007: Normal registration fee of € 150.

## Paper submission

Papers can be submitted through the conference website. Instructions will follow.

## Special Issue

Papers containing strong computational statistical or substantive data-analytic elements will be considered for publication in a special peer-reviewed, or regular, issue of the journal *Computational Statistics & Data Analysis*.

[URL:www.elsevier.com/locate/csda](http://www.elsevier.com/locate/csda)

## Scientific Program Committee

[Jörg Blasius](#), University of Bonn,  
[Michael Greenacre](#), Pompeu Fabra University,  
Patrick J.F. Groenen, Erasmus University  
Rotterdam,  
Michel van de Velden, Erasmus University  
Rotterdam.

## Local Organizing Committee

Patrick J.F. Groenen,  
Michel van de Velden,  
Joost van Rosmalen,  
Elli Hoek van Dijke,  
Martijn Kagie,  
Georgi Nalbantov.

## IASC International Conference on Statistics for Data Mining, Learning and Knowledge Extraction – Aveiro, Portugal, 2007



The International Conference on Statistics for Data Mining, Learning and Knowledge Extraction, a Satellite Conference of the 56<sup>th</sup> Session of the ISI, will take place in Aveiro, Portugal, on the 30th, 31st August and 1st September 2007. The purpose of this meeting is to foster the interaction of researchers in the interface between computational statistics, data mining, knowledge discovery and statistical learning. The meeting will consist of invited lectures, specialized sessions, contributed papers, posters and software demonstrations. Further information may be found in the conference webpage, <http://www.mat.ua.pt/iasc07>.

## COMPSTAT 2008



The 18th COMPSTAT symposium will take place in Porto, Portugal, on August 24th-29th 2008. The conference, sponsored by the European Regional Section of the IASC, is locally organized by the Faculty of Economics of the University of Porto ([www.fep.up.pt](http://www.fep.up.pt)).

Further information will be posted on the conference website, [www.fep.up.pt/compstat08/](http://www.fep.up.pt/compstat08/).

## The 4<sup>th</sup> World Conference on Computational Statistics & Data Analysis of the International Association for Statistical Computing (IASC) and the 6<sup>th</sup> Conference of Asian Regional Section of IASC



The Joint Meeting of 4th World Conference of the IASC and 6th Conference of the Asian Regional Section of the IASC on Computational Statistics & Data Analysis (IASC2008) will be held at Pacifico Yokohama in Japan on 5th-8th December 2008.

The conference aims at bringing together researchers and practitioners to discuss recent developments in computational methods, methodology and applications in statistics. It is associated with the Computational Statistics & Data Analysis (CSDA), the official journal of the IASC. E-mail: [iasc2008@ism.ac.jp](mailto:iasc2008@ism.ac.jp); website: <http://www.iasc-ars.org/IASC2008/>

## IASC-ERS Summer School on Data Mining and Statistical Learning Techniques, 2007

An IASC-ERS Summer School on Data Mining and Statistical Learning Techniques will take place in the Island of Procida, Italy from September the 3rd until September the 7th, 2007. The IASC-ERS school is intended to achieve postgraduate training in special areas of statistics. The participants are expected to have a good background in statistics at the Ph.D level, although not necessarily oriented to the subject of the course. Also professionals working in industry interested in Data Mining are invited to participate. The participation of young researchers is favoured through scholarships which will be limited and granted according to strictly criteria.

For further information please contact Prof. Rosaria Lombardo, [rosaria.lombardo@unina2.it](mailto:rosaria.lombardo@unina2.it).

Paula Brito

## New journal on data analysis and classification

During the GfKl conference 2006 in Berlin and the IFCS conference 2006 in Ljubljana Springer-Verlag has announced the publication of the new journal

### *Advances in Data Analysis and Classification (ADAC)*

beginning in Spring 2007. The journal arose from joint activities of members of CLADAG, GfKl, and JCS; its Editorial Board includes colleagues from various IFCS Member Societies. The journal deals with a broad range of data analysis methods, with classification and clustering, and knowledge extraction tools. Some emphasis is laid to new applications and recent application fields. Detailed information can be obtained from the journal's website at [www.springer.com/journal/11634](http://www.springer.com/journal/11634) (see also the former IFCS Newsletter).

Given that IFCS supports the journal *ADAC* as well as the *Journal of Classification*, Springer will offer *ADAC* to all individual members of IFCS Member Societies at an annual subscription price of maximally 42 Euro (plus VAT that is country-dependent). This is a considerable reduction in comparison with the price of 148 Euro for institutions and other individuals. Subscriptions must be collected by a Society member (typically: the treasurer) who thereby confirms membership. Lower prices are even possible for societies with a compulsory subscription rule. For details please ask one of the editors.

**Hans-Hermann Bock**  
**Wolfgang Gaul**  
**Akinori Okada**  
**Maurizio Vichi**

### **Short book review**

B. Le Roux, H. Rouanet *Geometric Data Analysis*

A little more than 30 years from the publication of the two volumes by J.P. Benzécri *L'Analyse des données*, a milestone

and an obliged reference for anyone who wants to approach the issues and the questions related to the exploration and the description of the phenomena we are dealing with daily, Brigitte Le Roux and Henry Rouanet publish a volume that, already in the title, enhances the attempt to value the geometric approach, that is the possibility to represent the data as clouds of points in a multidimensional Euclidean space and, after all, the possibility to visualize relations that “*may be described by numbers but are not reducible to numbers*”.

The text runs through the fundamental stages of Multidimensional Data Analysis, facing the methods that sanctioned its success, in particular Principal Component Analysis (PCA) and Correspondence Analysis, illustrating some of their methodological principles and supplying precious instruments of interpretation. On the other hand, the authors do not evade the spiny issues that have characterized the debate in the last twenty years between “English-speaking” and “French” school on the possibility to use information that contributes to the definition of the projection space even not being part of the “characterizing” set of variables (Structured Data Analysis), or on the possibility to generalize the conclusions to something wider than the specific set of observed data (Inductive Data Analysis).

The text starts with a methodological overview, in which the fundamental principles of Geometric data Analysis are reminded, i.e. the possibility to represent multivariate data sets as clouds of points in proper metric spaces, the importance of a mathematical formalism that justifies and strengthens the proposed procedures and, finally, what we could call a description-oriented approach.

Chapters 2 through 5 propose the methodological bases of the most important methods in Multidimensional Data Analysis, that is Correspondence Analysis, Multiple Correspondence Analysis and Principal Component Analysis. The methods are explained with methodological rigour and, probably, the graphical aspects of their more formal part partially penalizes the text itself,

that turns out too much “crowded” of formulas and, just for that, not always easy to read. On the other hand, the examples, the exercises, the comments and the aids to the interpretation that the authors show in the second part of each chapter, offer stimulating cues of deepening not only for those who are approaching these subjects for the first time but also for those who are more regular users of the presented methods.

Chapters 6 through 8 face fundamental and still open issues that, respectively, regard the Data Structure, the problems of stability and, as a consequence, the inferential aspects. Still in this section, the links with the French School are confirmed by, as an example, the re-evaluation of the individual-points and the role that the projection of clouds of points, even though “anonymous”, may have in the understanding of the observed phenomenon.

Chapter 9 is a precious “book in the book”. Three study cases are clearly presented and deeply analyzed. The reader can “see” what he has just learnt, from a methodological point of view, in previous chapters.

The text ends with an always useful chapter of call-backs on the basis of Matrix Algebra and Vector Spaces.

*Geometric Data Analysis* therefore appears as an important reference for those who want to approach this fundamental chapter of statistics but also for those who already know its properties and qualities and want to explore some of the possible deeper insights. A less “severe” graphical layout would be probably more favourable for an easier reading.

Marco Gherghi,  
Department of Mathematics and Statistics,  
University of Naples “Federico II” – Italy

**Contributions for the coming issue of the IFCS newsletter can be sent to:**

andre.hardy@fundp.ac.be

Text files are by preference in ascii or word, with pc format. For graphical materials, by preference GIF is used.

Editor:

André HARDY

Department of Mathematics

University of Namur

8, Rempart de la Vierge

B – 5000 Namur

Belgium

Tel : +32 81 72 49 05

Fax : +32 81 72 49 14