

40 Years

Gesellschaft für Klassifikation
– From Numerical Taxonomy to Big Data –

Hans-Hermann Bock

1. The origins
2. Activities - Conferences - Topics
3. Publication activities
4. Persons and working groups
5. Aspects for the future

GfKI Jubilee Colloquium, Karlsruhe, February 24, 2017

GfKI was founded on February 12, 1977 at Frankfurt Main Station

8 founding members:

Dr. Ingetraut Dahlberg

Dr. Robert Fugmann

Dr. Hans-Hermann Bock

Dipl.-Volksw. Otto Gekeler

Prof. Dr. Martin Scheele

Dr. Hans-Georg Schön

.....

Philosophy, documentation (DGD)

Chemical databases (Hoechst)

Mathematics, statistics (TU Hannover)

Commodity classification (BWL, AEG)

Biological taxonomy/documentation

Librarian (UB Münster)



Classification in 1960-1977

- **Broad interest in classification methods**

Classification and indexing for libraries
(Dewey Decim. Classif., Thesauri)

Classification systems in documentation
e.g., chemical documentation

Universal classification (of all entities ...)

Product and commodity classification

Patent classification

Index Medicus

Biological taxonomy and systematics

Mathematical methods, classification
rules, clustering algorithms

Analyse des données

- **Societies:**

Gesellschaft für Dokumentation (DGD)

Gesellschaft für Information und
Dokumentation (GID)

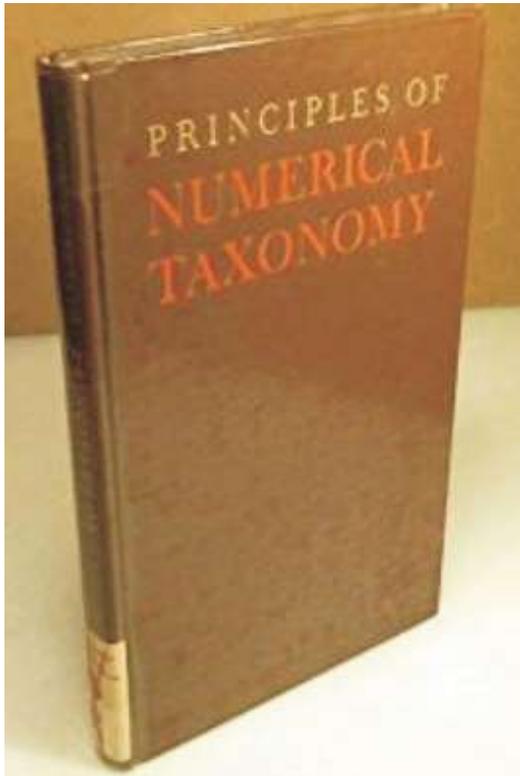
Fédération Intern. de Documentation (FID)
Committee 'Classification Research'

International Society for Commodity
Classification and Technology (IGWT)

Infoterm (Vienna)

Systematics Association (USA, UK)

Classification Society (USA, UK)



R.R. Sokal, P.H.A. Sneath
**Principles of
Numerical Taxonomy**
1965

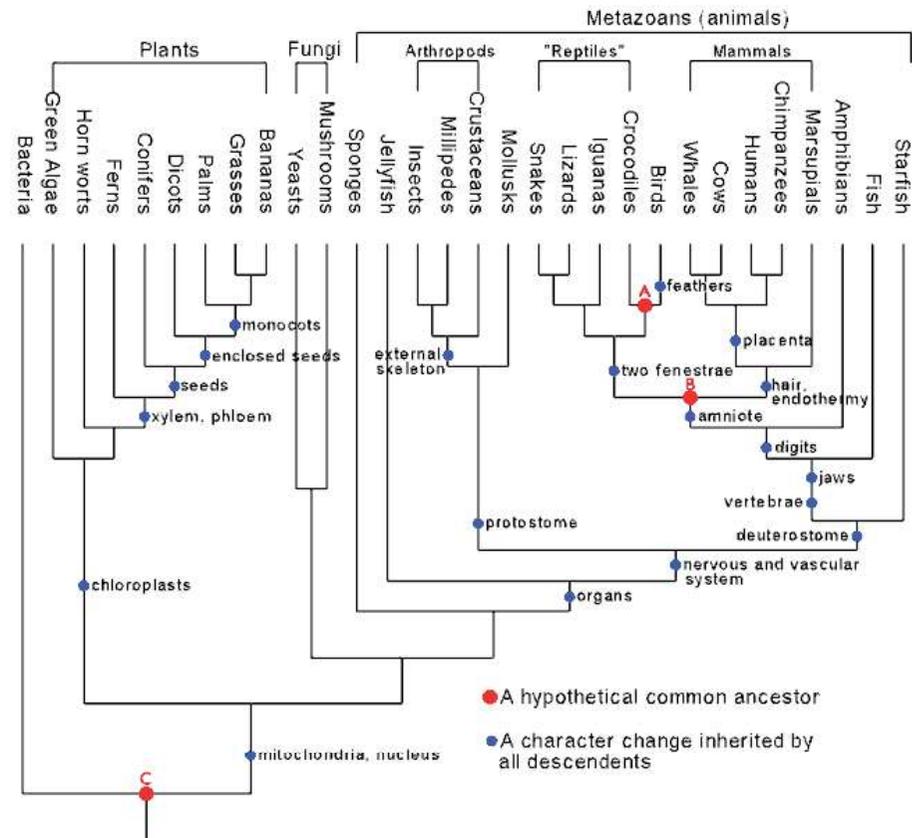
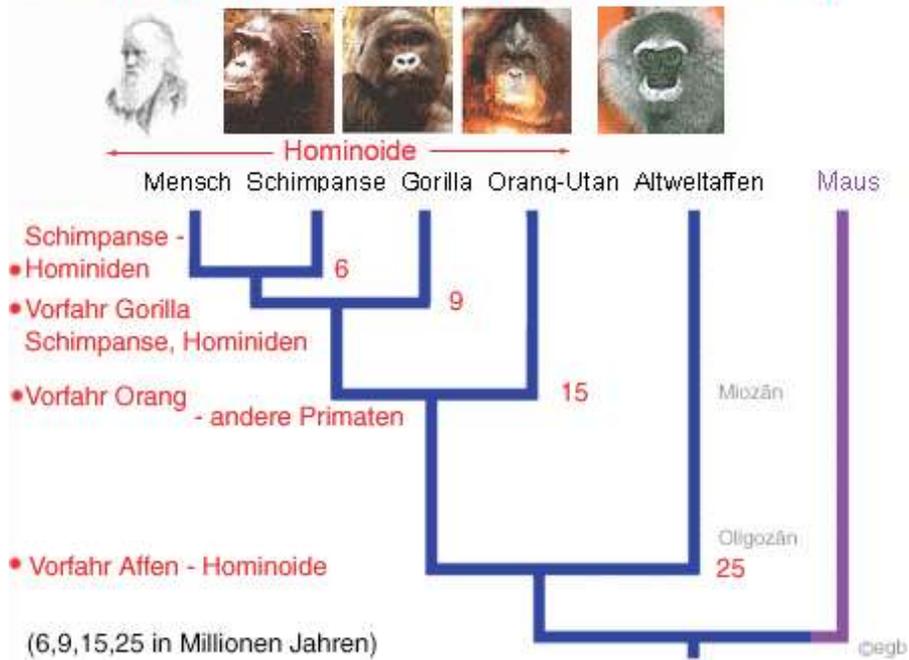


I. Dahlberg
**Grundlagen universaler
Wissensordnung**
1974

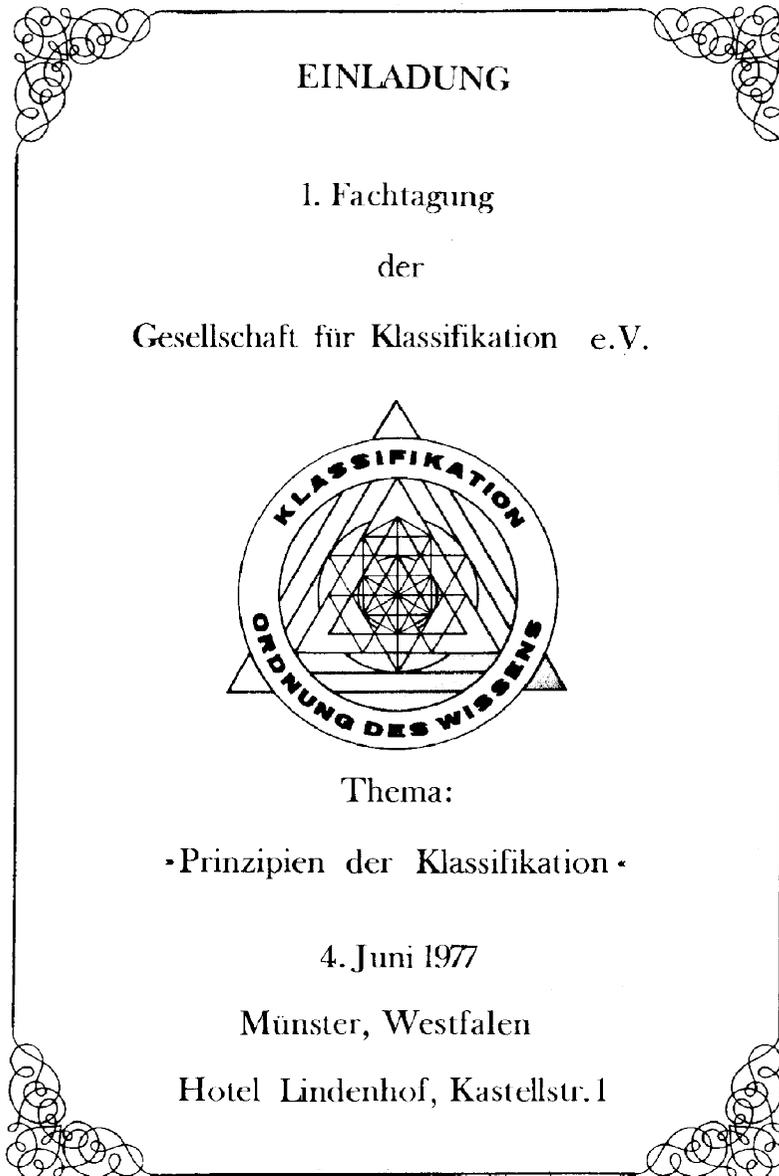


H.-H. Bock
**Automatische
Klassifikation**
1974

Cladogramm der Hominiden (DNA-Hybridisierung)



2. Activities - conferences - topics

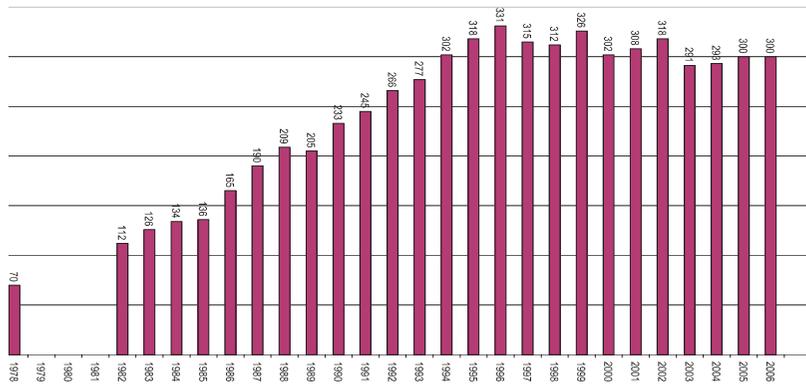


First conference in Münster, June 4, 1977

with 44 participants and 9 papers:

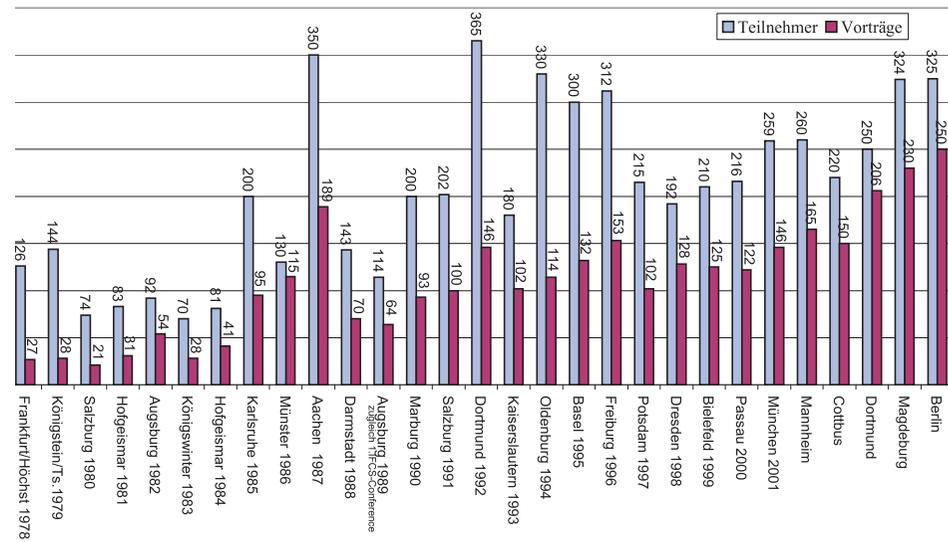
- Linguistic bases of scientific conceptualization (Gipper)
- Conceptual bases of classification (Dahlberg)
- A dichotomous classification of 'definitions' (Nacke)
- Usage of the term 'concept' (Nacke)
- Reproduction of concept relations by a classification
- Fundamentals of numerical classification (Bock)
- Bases of ordination and seriation (Ihm)
- Classification principles in libraries (Lohse)

Number of members of GfKI



The conferences: # participants/lectures

Jahrestagungen der Gesellschaft für Klassifikation



From the 'conceptual period' to 'statistical data analysis':

1977 - 1989

from about 1984

First topics in GfKI:



The new topics:

Order and concept theory

Philosophical bases

Universal classification

Linguistic methods

Documentation

Thesaurus construction

Automatic indexing

Library classification

Numerical classification

Exploratory data analysis

Discriminant analysis, pattern recognition

Clustering models and algorithms

Multidimensional scaling, visualization

Correspondence analysis

Formal concept analysis (lattice theory)

Preference structures

Many applications of numerical methods

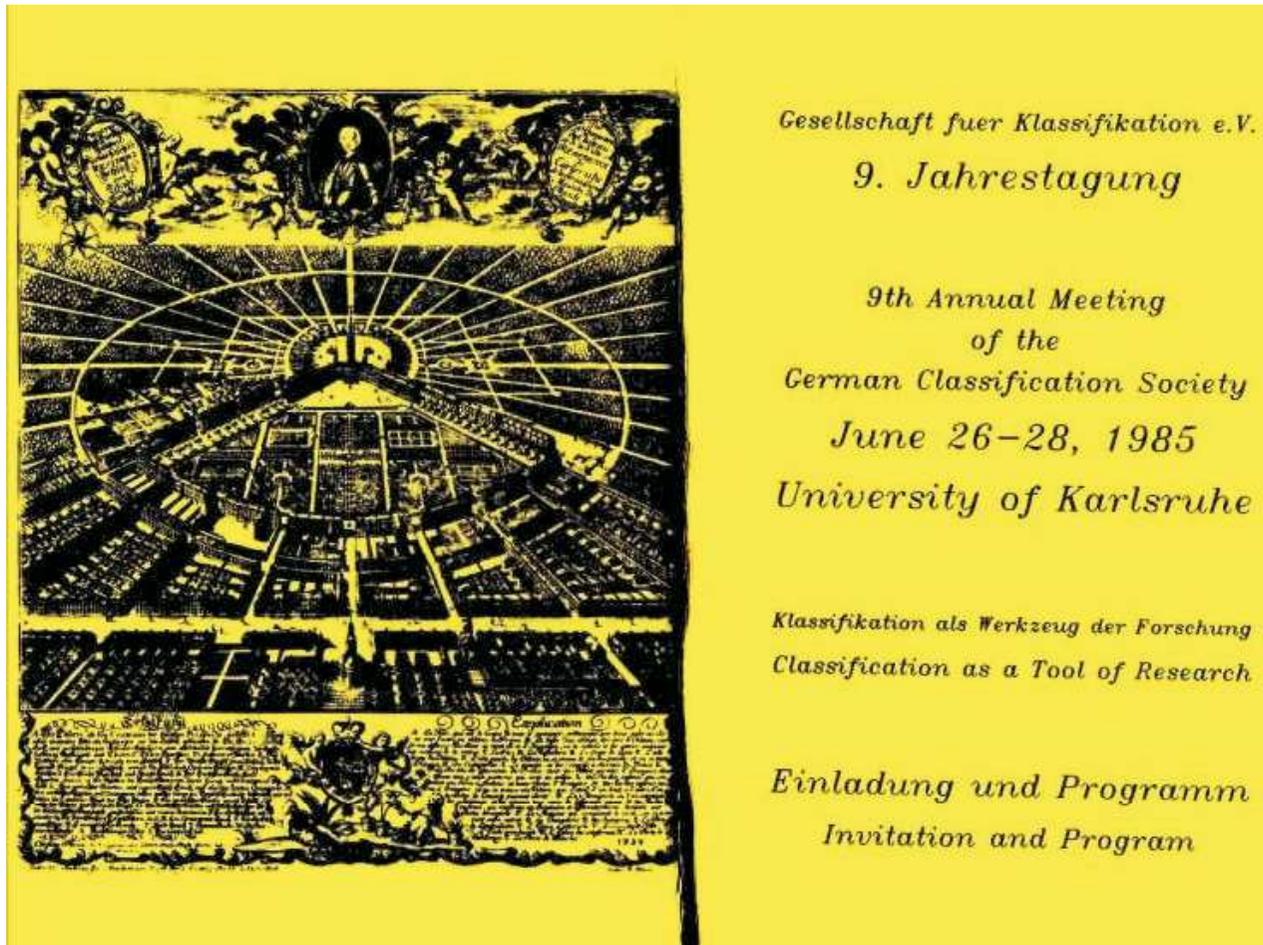


1989: New Society splits off:

International Society for Knowledge Organisation (ISKO)

International relationships 1

- 9th GfKI conference in Karlsruhe (1985, W. Gaul)



with many internationally renowned experts:

Arabie, Carroll, Critchley, De Leeuw, Hubert, Murtagh, De Soete, Wishart

July 1985, Cambridge, UK:
 Founding the

International Federation of Classification Societies (IFCS)

We, the approved delegates of the following Societies or Groups:

- British Classification Society (BCS)
- Classification Society of North America (CSNA)
- Gesellschaft für Klassifikation e.V. (GfKI)
- Japanese Classification Society (JCS)
- Société Francophone de Classification (SFC)
- Italian Statistical Society (SIS)

declare the formation of the

INTERNATIONAL FEDERATION OF CLASSIFICATION SOCIETIES

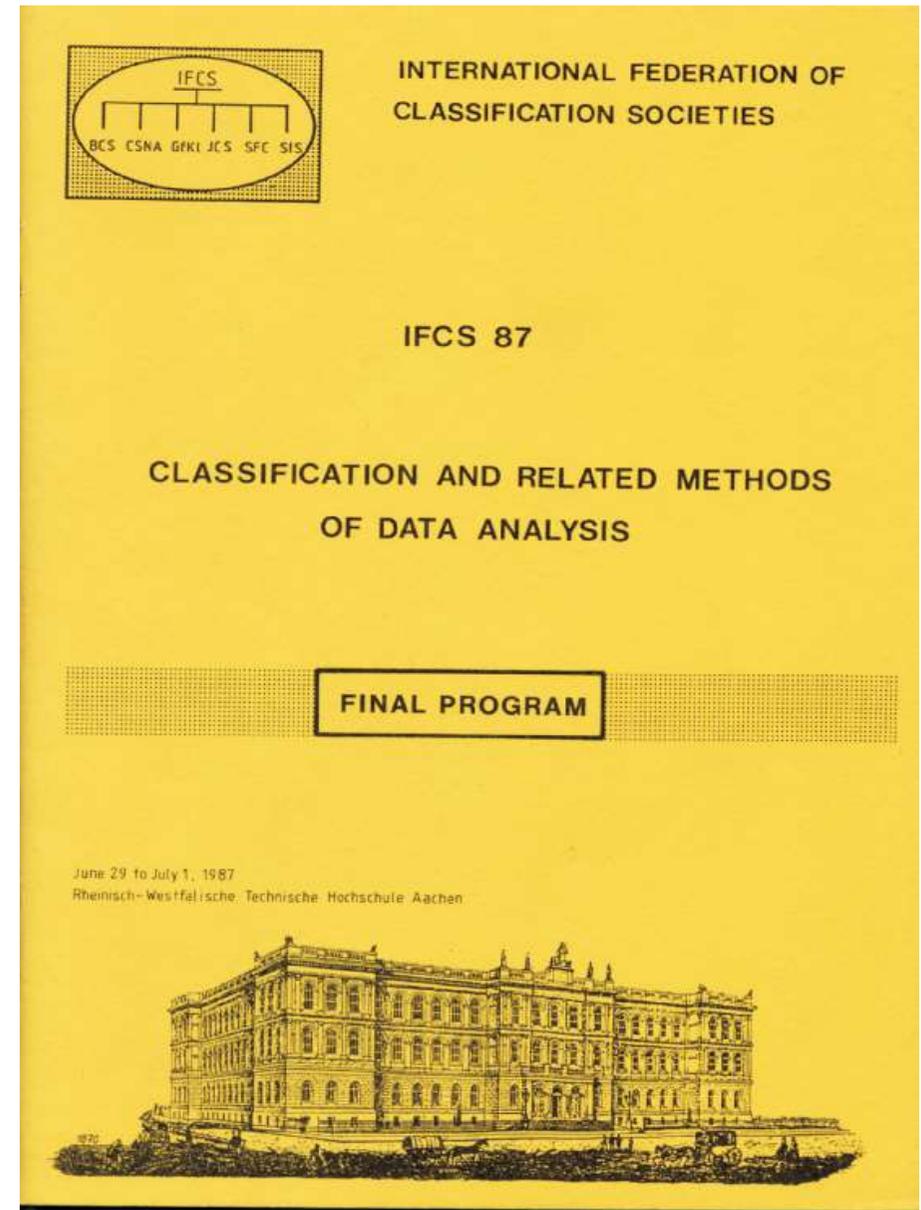
the founding members of which are as listed above. The provisional constitution is given as an annexe; a final constitution will be adopted by unanimous agreement of the founding Council.

Queens' College, Cambridge, U.K.

July 4th, 1985

J.C. Gower	<i>J.C. Gower</i>	A.J. Boyce	<i>Anthony Boyce</i> (BCS)
J.D. Carroll	<i>J.D. Carroll</i>	R.R. Sokal	<i>R.R. Sokal</i> (CSNA)
H-H. Bock	<i>H.H. Bock</i>	P. Ihm	<i>Peter Ihm</i> (GfKI)
C. Hayashi	<i>C. Hayashi</i>	K. Yajima	<i>Keiji Yajima</i> (JCS)
M. Jambu	<i>M. Jambu</i>	C. Perruchet	<i>C. Perruchet</i> (SFC)
N. Lauro	<i>N. Lauro</i>	A. Mineo	<i>A. Mineo</i> (SIS)

July 1987, Aachen, Germany:
 1st Conference of the IFCS



Following the methodological trends at conferences:

GfKI Conference 1998 (Dresden):

Statistical data analysis

Classification and clustering
Formal concept analysis
Learning algorithms

Internet and information technology
Usage of new media and internet
Information systems
Classification for multimedia, WWW
Navigation and classif. on the internet
Virtual universities

Applications

Marketing, market research
Finance, capital, risk management

Bioinformatics, Genomics,...

Environment
Health sciences

Archaeology

GfKI Conference 2007 (Freiburg)

Statistical data analysis

- Multidimensional data analysis
- Symbolic data analysis
- Data visualization
- Time series and panel data
- Image and signal analysis

Classification

- Supervised classif., pattern recognition
- Cluster analysis
- Similarity structures
- Mixture analysis, latent class models
- Classification and regression

Data mining

- Analyzing data streams
- Text mining, web mining
- Semantic web
- Statistical relational learning
- Knowledge representation and discovery

Marketing and management science

- Banking and finance
- Data analysis in retailing
- Business intelligence and personalization

Various applications

- Biostatistics and bioinformatics
- Medical and health sciences
- Statistical musicology
- Linguistics
- Archaeology and archaeometry
- Library classification, automatic indexing

International relationships 2

Since 2013:

- GfKI conferences are organized as "European Conference on Data Analysis" (ECDA)
- related to the foundation of a "European Association for Data Science" (EuADS)
- incorporated in Luxembourg

ECDA	Place	Cooperating societies
2013:	University of Luxembourg	French Classification Society SFC
2014:	Jacobs-Universität Bremen	Italian CLADAG, Dutch VOC, Polish SKAD, IASC
2015:	Univ. of Essex, Colchester	SKAD and British Classification Society BCS
GfKI-2016:	DAGStat, Göttingen	GfKI, Biometric Society/DR, ...
2016:	European Data Science Conf.	EuADS in Luxembourg-Grund
2017:	Wroclaw, Poland	Polish SKAD, European EuADS, Italian CLADAG, British BCS

39th Annual Conference of GfKI:

European Conference on Data Analysis - ECDA-2015

Colchester/UK

Data analysis and statistics (8)

- Supervised classification (2)
- Clustering (3)
- Optimization methods (1)
- Data analytics (1)
- Crowdsourcing (1)
- Machine learning
and knowledge discovery (4)
- Big data (1)
- Math. foundations of Data Science (2)

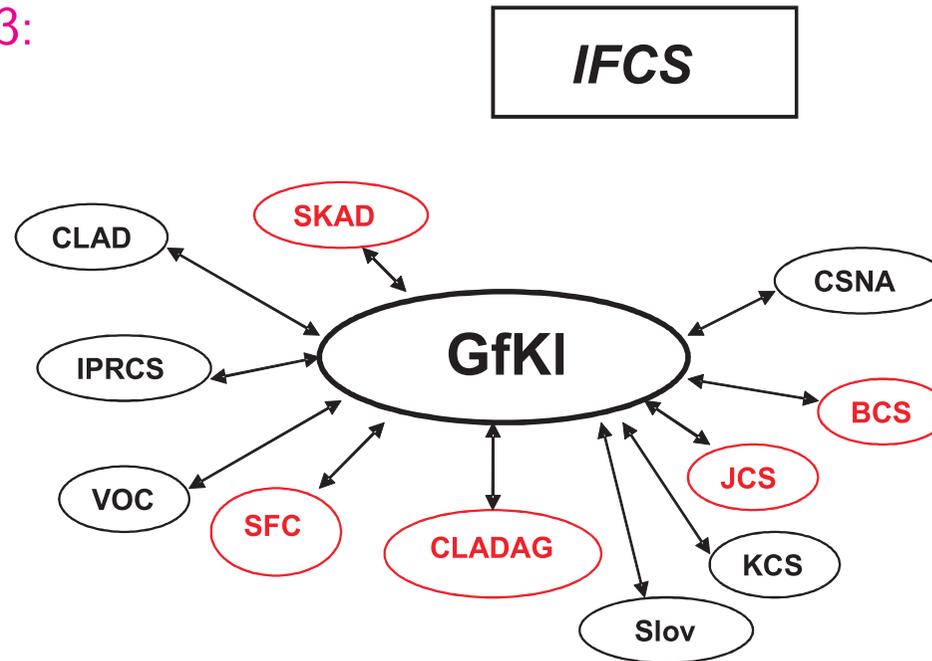
Applications

- Economics and finance (3)
- Marketing (2)
- Engineering, logistics (1)
- Biostatistics, bioinformatics, life sciences (1)
- Digital humanities, social sciences (2)
- Education (1)
- Archaeology and geosciences (1)
- Musicology (1)

Library classification, automatic indexing

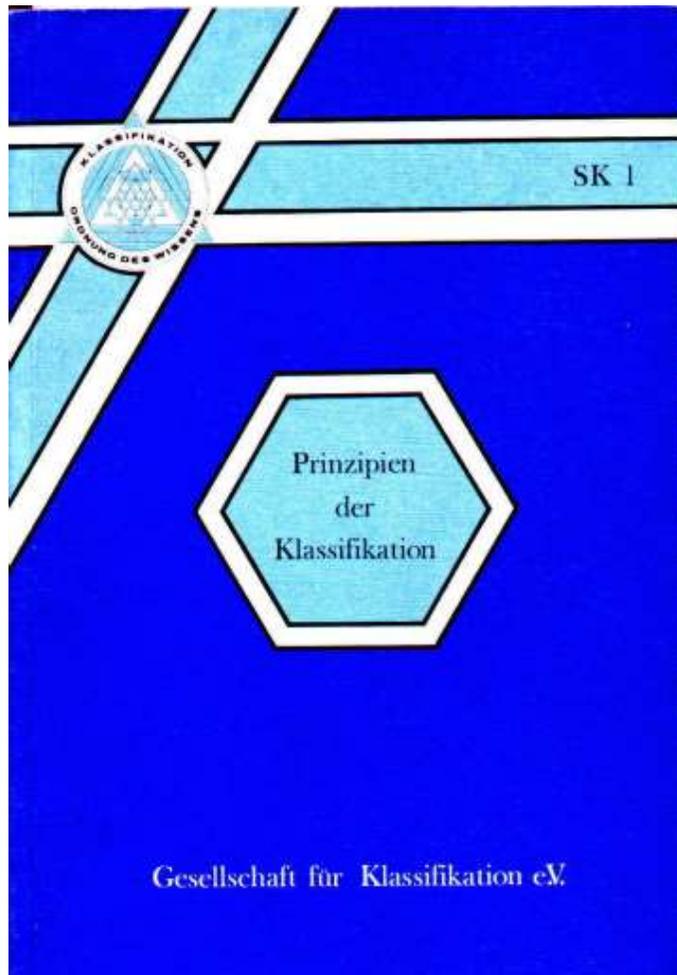
International relationships 3:

Network of IFCS Societies:

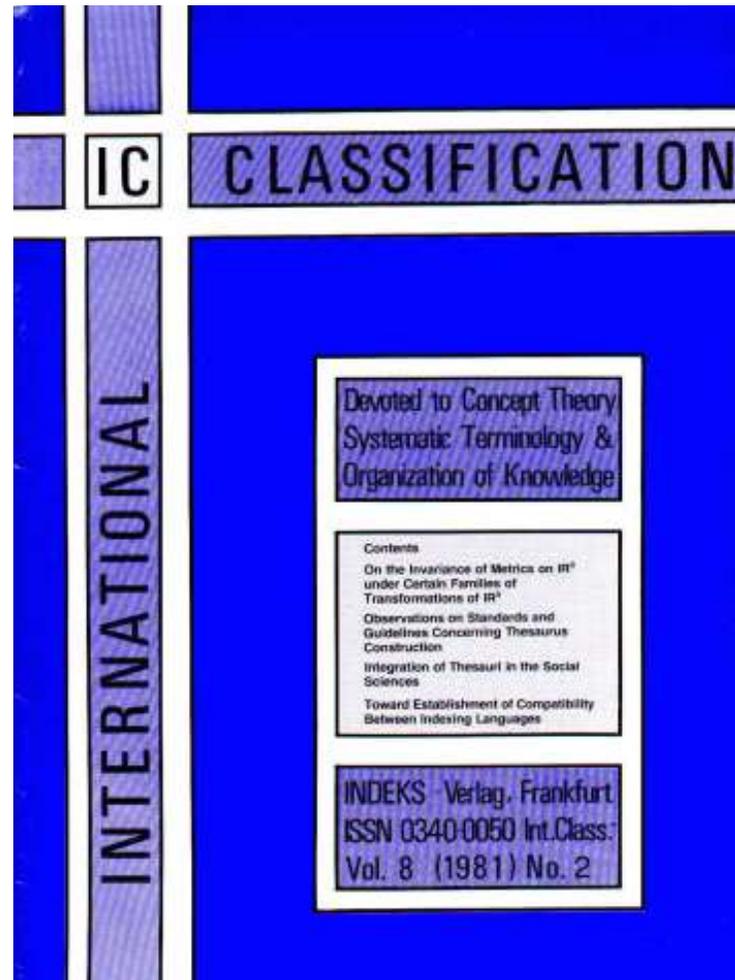


- Organization of international IFCS conferences:
 - 1987 Aachen, 2009 Dresden, 2011 Frankfurt
- Biennial bilateral 'Symposia on Data Analysis and its Applications'
 - GPSDA with Polish Classification Society (SKAD)
Aachen 2009, Krakau 2011, Dresden 2013
 - GJSDA with Japanese Classification Society (SKAD)
Tokyo 2005, Berlin 2006, Karlsruhe 2010, Kyoto 2012, Reicensburg 2016
- Special joint conferences, e.g.:
 - with Italian CLADAG: Firenze 2010

3. Publications activities

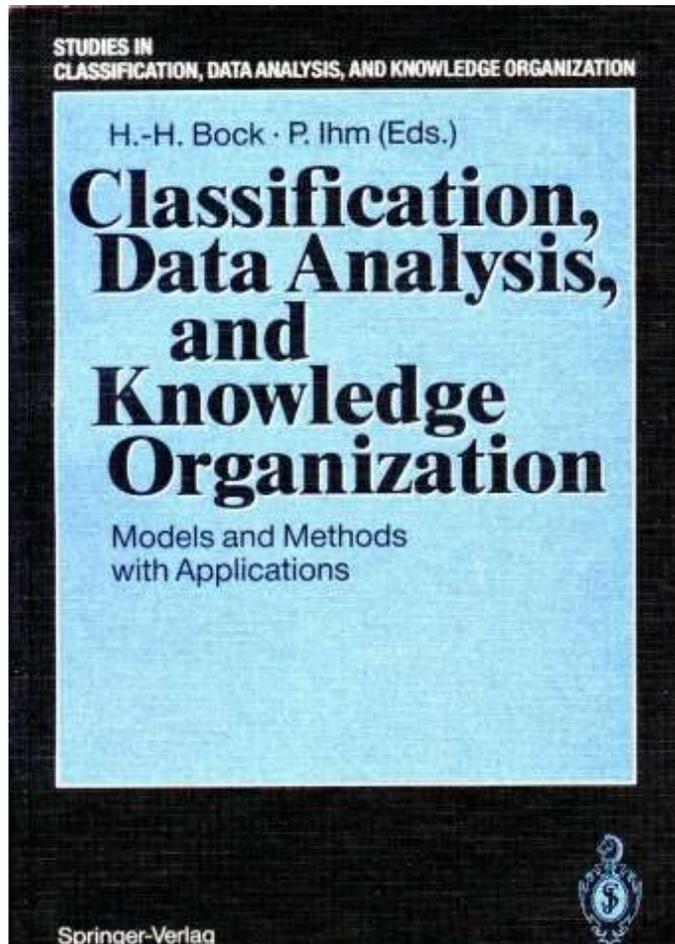


Studien zur Klassifikation
1 (1977) - 19 (1989)



International Classification
1 (1974) - 19 (1992)

Studies in Classification, Data
Analysis and Knowledge Organization
Since 1989, thus far: 50 volumes



Advances in Data Analysis
and Classification (ADAC)
1 (2007) - 10 (2017)



Since 2013/14:

Archives of Data Science

- Open Access publications of GfKI -

Editorial Board: A. Geyer-Schulz, E. Hüllermeier, H. Kestler

KIT - Scientific Publishing

www.ArchivesforDataScience.org

4. Persons and working groups



Ingetraut Dahlberg
1977 - 1986



Hans-Hermann Bock
1986 - 1995

The presidents of GfKI



Wolfgang Gaul
1995 - 1998



Otto Opitz
1998-2004



Claus Weihs
2004-2013



Berthold Lausen
since 2013

Working Groups:

Datenanalyse und Numerische Klassifikation (AG-DANK)

Since 1978



G. Ritter



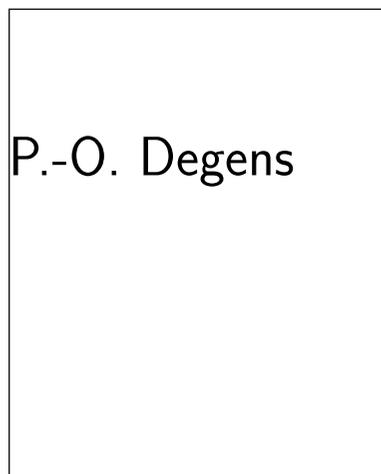
Chr. Hennig



Hans-J. Mucha

Datenanalyse und Klassifikation in den Biowissenschaften (AG-BT)

Since 1987: Biologische Taxonomie; P.O. Degens, W. Ludwig, B. Lausen, H. Kestler



P.O. Degens



B. Lausen



H. Kestler

Computeranwendungen und Quantitative Methoden in der Archäologie (AG-ARCH, CAA)

since 1990, chaired by

P. Ihm †, A. Zimmermann, I. Herzog, K. Lambers



Bibliotheken, Bibliotheksklassifikation (AG-BIB)

since 1979: H. Düvel,....., F. Scholze



Dezimalklassifikationen (AG-DK)

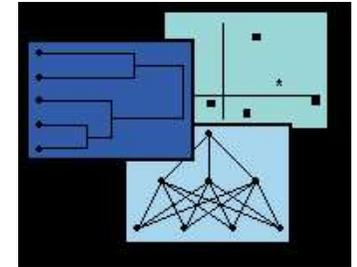
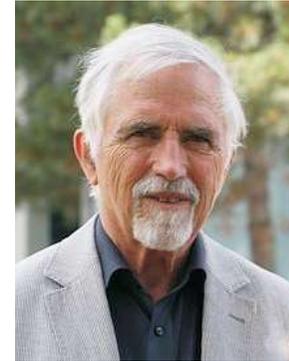
since 1992, chaired by

G. Knell, A. Weiße, H.-J. Hermes, B. Lorenz



Specific topics:

Formale Begriffsanalyse (Rudolf Wille †)



Linguistik, Dialektometrie (Hans Goebel)

Economics (Gaul, Opitz, Decker, Baier,...)



Musicology (Claus Weihs)

The current Board of GfKI



5. Aspects for the future: Data Science and Big Data

Major tasks for GfKI

- Attracting young scientists
- New problem fields, methods and applications
- Data Science
- External profiling

Big Data

- Role of experimental design
- Differing types of validity of results
- Ethical guidelines