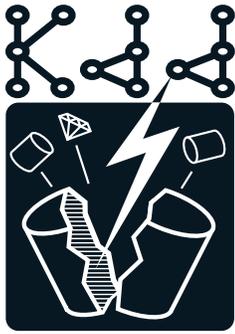




The Third International Conference on Knowledge Discovery and Data Mining

August 14–17 1997, Newport Beach, California

Sponsored by the American Association for Artificial Intelligence



KDD-97: A Preview

The rapid growth of data and information has created a need and an opportunity for extracting knowledge from databases, and both researchers and application developers have been responding to that need. Knowledge discovery in databases (KDD), also referred to as data mining, is an area of common interest to researchers in machine discovery, statistics, databases, knowledge acquisition, machine learning, data visualization, high performance computing, and knowledge-based systems. KDD applications have been developed for astronomy, biology, finance, insurance, marketing, medicine, and many other fields.

The Third International Conference on Knowledge Discovery and Data Mining (KDD-97) will follow up the success of KDD-95 and KDD-96 by bringing together researchers and application developers from different areas focusing on unifying themes.

Keynote Address: From Large to Huge. A Statistician's Reactions to KDD and DM

Peter Huber, *Universität Bayreuth, Germany*

The statistics and AI communities are confronted by the same challenge, the onslaught of ever larger data collections, but the two communities have reacted independently and differently. What could they learn from each other if they looked over the fence? What is amiss on either side?

KDD-97 Organization

General Conference Chair

Ramasamy Uthurusamy, *General Motors Corporation*

Program Cochairs

David Heckerman, *Microsoft Research*
Heikki Mannila, *University of Helsinki, Finland*
Daryl Pregibon, *AT&T Laboratories*

Publicity Chair

Paul Stolorz, *Jet Propulsion Laboratory*

Tutorial Chair

Padhraic Smyth, *University of California, Irvine*

Demo and Poster Sessions Chair

Tej Anand, *NCR Corporation*

Awards Chair

Gregory Piatetsky-Shapiro, *GTE Laboratories*

Program Committee Members

Helena Ahonen, *University of Helsinki*
Tej Anand, *NCR*
Ron Brachman, *AT&T Laboratories*
Carla Brodley, *Purdue University*
Dan Carr, *George Mason University*
Peter Cheeseman,
NASA AMES Research Center
David Cheung, *University of Hong Kong*
Wesley Chu, *University of California at Los Angeles*

Gregory Cooper, *University of Pittsburgh*
Robert Cowell, *City University, UK*
Bruce Croft, *University of Massachusetts at Amherst*
William Eddy, *Carnegie Mellon University*
Charles Elkan, *University of California at San Diego*
Usama Fayyad, *Microsoft Research*
Ronen Feldman, *Bar-Ilan University, Israel*
Jerry Friedman, *Stanford University*
Dan Geiger, *Technion, Israel*
Clark Glymour, *Carnegie-Mellon University*
Moises Goldszmidt, *SRI International*
Georges Grinstein, *University of Massachusetts, Lowell*
Jiawei Han, *Simon Fraser University, Canada*
David Hand, *Open University, UK*
Trevor Hastie, *Stanford University*
David Heckerman, *Microsoft Corporation*
Haym Hirsh, *Rutgers University*
Jim Hodges, *University of Minnesota*
Se June Hong, *IBM T. J. Watson Research Center*
Tomasz Imielinski, *Rutgers University*
Yannis Ioannidis, *University of Wisconsin*
Lawrence D. Jackel, *AT&T Laboratories*
David Jensen, *University of Massachusetts*
Michael Jordan, *Massachusetts Institute of Technology*
Daniel Keim, *University of Munich*
Willi Kloesgen, *GMD, Germany*
Ronny Kohavi, *Silicon Graphics*
David D. Lewis, *AT&T Laboratories – Research*
David Madigan, *University of Washington*
Heikki Mannila, *University of Helsinki*
Brij Masand, *GTE Laboratories*
Gary McDonald, *General Motors Research*

Eric Mjolsness, *University of California at San Diego*
Sally Morton, *Rand Corporation*
Richard Muntz, *University of California at Los Angeles*
Raymond Ng, *University of British Columbia*
Steve Omohundro, *NEC Research*
Gregory Piatetsky-Shapiro,
GTE Laboratories
Daryl Pregibon, *Bell Laboratories*
Raghu Ramakrishnan, *University of Wisconsin, Madison*
Patricia Riddle, *Boeing Computer Services*
Ted Senator, *NASD Regulation Inc.*
Jude Shavlik, *University of Wisconsin at Madison*
Wei-Min Shen, *University of Southern California*
Arno Siebes, *CWI, Netherlands*
Avi Silberschatz, *Bell Laboratories*
Evangelos Simoudis, *IBM Almaden Research Center*
Andrzej Skowron, *University of Warsaw*
Padhraic Smyth, *University of California at Irvine*
Ramakrishnan Srikant, *IBM Almaden Research Center*
John Stasko, *Georgia Institute of Technology*
Salvatore J. Stolfo, *Columbia University*
Paul Stolorz, *Jet Propulsion Laboratory*
Hanna Toivonen, *University of Helsinki*
Alexander Tuzhilin, *New York University, Stern School*
Ramasamy Uthurusamy, *General Motors R&D Center*
Graham Wills, *Bell Laboratories*
David Wolpert, *IBM Almaden Research Center*
Jan Zytkow, *Wichita State University*

KDD-97 Tutorial Program

All tutorials will be presented on Thursday, August 14, 1997. The times listed below are tentative. Admission to the tutorials is included in your conference registration fee. Registrants can attend up to four consecutive tutorials, including four tutorial syllabi.

TIME	SESSION 1	SESSION 2
8:00 to 10:00 AM	T1 – Fayyad and Simoudis (<i>single session</i>)	
10:30 AM to 12:30 PM	T2 – Hand	T3 – Feldman
1:30 to 3:30 PM	T4 – Swayne and Cook	T5 – Chaudhuri and Dayal
4:00 to 6:00 PM	T6 – Keim	T7 – DuMouchel

Tutorial 1: 8:00–10:00 AM

Data Mining and KDD: An Overview

Usama Fayyad, Microsoft Research and Evangelos Simoudis, IBM

We present a basic tutorial of this new and emerging area and emphasize relations to constituent communities, including statistics, databases, pattern recognition, learning, and visualization. The tutorial provides a basic overview of the KDD process for extracting knowledge from databases and covers the basics of each step in the process including: data warehousing, selection and cleaning, data transformation, data mining, evaluation, and visualization. We also cover a sampling of successful applications and outline challenges and issues to be addressed.

Usama Fayyad is a senior researcher at Microsoft Research, the Decision Theory & Adaptive Systems Group. His research interests include knowledge discovery in large databases, data mining, machine learning, statistical pattern recognition, and clustering. After receiving a Ph.D. degree in 1991, he joined the Jet Propulsion Laboratory (JPL), California Institute of Technology (until 1996). At JPL, he headed the Machine Learning Systems Group where he developed data mining systems for analysis of large scientific databases.

Evangelos Simoudis is Vice President, Global Business Intelligence Solutions – IBM North America, where he is responsible for the development and deployment of data mining and decision support solutions to IBM's customers worldwide. Simoudis received a B.A. in physics from Grinnell College, a B.S. in electrical engineering from California Institute of Technology, an M.S. in computer science from the University of Oregon, and a Ph.D. in computer science from Brandeis University.

Tutorial 2: 10:30 AM–12:30 PM

Modeling Data and Discovering Knowledge

David Hand, Open University, UK

Our aim is to extract knowledge from large bodies of data. The size of these bodies mean that we cannot do it unaided, but must use fast computers, applying sophisticated statistical tools. Attempts to automate the process of knowledge extraction date from at least the early 1980s, with the work on statistical expert systems. We examine this work, noting its successes and failures and especially what researchers in data mining and knowledge discovery can learn from those efforts. We examine what data are, what information is, and what knowledge is. We contrast modeling with discovery, especially in the context of large data sets. We examine high level modeling issues, such as overfitting, generalisability, overmodeling, and model evaluation. And we examine high level exploration issues such as the discovery of accidental artifacts. The confluence of computing and statistics in some areas provides a nice backdrop against which to examine these issues, and we briefly discuss neural networks and classification trees from these two perspectives.

David Hand is a professor of statistics at the Open University. His research interests include the foundations of statistics, statistical computing, and multivariate statistics, the latter especially as applied to classification problems. His applications interests include medicine, finance, and psychology. He is editor-in-chief of *Statistics and Computing* and has published fourteen books, the most recent of which is *Construction and Assessment of Classification Rules*, (Wiley, January 1997).

Tutorial 3: 10:30 AM–12:30 PM

Text Mining— Theory and Practice

Ronen Feldman, Bar-Ilan University, Israel

Knowledge discovery in databases (KDD) focuses on the computerized exploration of large amounts of data and on the discovery of interesting patterns within them. While most work on KDD has been concerned with structured databases, there has been little work on handling the huge amount of information that is available only in unstructured textual form. In this tutorial we will present the general theory of text mining and will demonstrate several systems that use these principles to enable interactive exploration of large textual collections. We will describe generic techniques for text categorization and information extraction that are used by these systems. The systems that will be presented are KDT which is the system for knowledge discovery in texts; FACT, which discovers associations among keywords labeling the items in a collection of textual documents; and the Text Explorer, which is a system that provides a high level language for interactive exploration of textual collections. We will present a general architecture for text mining and will outline the algorithms and data structures behind the systems. We will give special emphasis to incremental algorithms and to efficient data structures.

Ronen Feldman is a lecturer at the Mathematics and Computer Science Department of Bar-Ilan University in Israel. He received his B.Sc. in math, physics and computer science from the Hebrew University, and his Ph.D. in computer science from Cornell University. His main research is in the area of machine learning and data mining. He is currently coordinating several research projects for developing dedicated text mining systems. These systems work on plain text collections and on the Internet.

Exploratory Data Analysis Using Interactive Dynamic Graphics

Deborah Swayne, Bell Communications Research and Diane Cook, Iowa State University

Researchers and software designers in the field of data mining are just beginning to make extensive use of graphical methods. Interactive dynamic data visualization has been explored in the field of statistics for over twenty years, and we propose that much of what has been learned in statistics is relevant for data mining. This class is an introduction to interactive data visualization as it is practiced as part of exploratory data analysis. The XGobi software, publicly available dynamic visualization software, will be used in the analysis of examples from biology, business, physics, engineering, and telecommunications. The examples will illustrate a set of general visualization principles which are embodied in specific methods such as brushing and identification of points in simple scatterplots, three dimensional rotations, rotations in higher dimensions such as the grand tour, and directed searches in higher dimensions for interesting two dimensional views using projection pursuit and manual control.

Deborah Swayne has worked at Bellcore since that company's inception in 1985, and is currently a member of the Statistics and Data Mining Research Group. Her research focuses on software methods for visualizing data. She is one of the authors of the XGobi software, originally developed at Bellcore. She has a B.A. in African linguistics from the University of Wisconsin at Madison, and a M.S. in statistics from Rutgers University.

Dianne Cook is an assistant professor in the Department of Statistics, Iowa State University. She received her PhD from Rutgers University in May 1993, and has conducted research into dynamic statistical graphics. Her interests include using these methods for understanding high-dimensional data, and adapting them for analyzing geographically referenced data with multiple measurements at each site.



Tutorial 5: 1:30–3:30 PM

OLAP and Data Warehousing

Surajit Chaudhuri, Microsoft Research and Umesh Dayal, Hewlett Packard Laboratories

On-line analytical processing (OLAP) and data warehousing technologies enable enterprises to gain competitive advantage by exploiting the ever-growing amount of data that is collected and stored in corporate databases and files for better and faster decision making. Over the past few years, these technologies have experienced explosive growth, both in the number of products and services offered, and in the extent of coverage in the trade press. Vendors (including database companies) are paying increasing attention to all aspects of decision support. The area opens up interesting research directions, with ties to past work in database systems, but with different assumptions and requirements. Only very recently, however, has the database research community started to understand and address some of these issues. This tutorial presents an overview of OLAP and data warehousing, and an in-depth study of selected aspects. An outline of the tutorial follows: 1) Introduction: definitions, evolution, differences from OLTP, architectures. 2) Models and tools: conceptual model for OLAP, front-end tools (e.g., multidimensional spreadsheets), database design (e.g., star and snowflake schema). 3) Database server technologies for decision support queries: specialized indexing techniques, specialized join and scan methods, data partitioning and use of parallelism, intel-

ligent processing of aggregates, complex query processing, extensions to SQL, ROLAP versus MOLAP. 4) Other services for OLAP/data warehousing: data cleaning, loading and refresh, tools for warehouse, system and process management, metadata management and the role of repository. 5) State of commercial practice. 6) Research issues.

The target audience is researchers and developers interested in learning about the concepts, products and the technical innovations in the area of decision support technologies.

Surajit Chaudhuri is a researcher in the Database Research Group of Microsoft Research. From 1992 to 1995, he was a member of the technical staff at Hewlett-Packard Laboratories, Palo Alto. He earned a B.Tech at the Indian Institute of Technology, Kharagpur and a Ph.D. at Stanford University. In addition to query processing and optimization, Chaudhuri is interested in the areas of data mining, database design and uses of databases for nontraditional applications.

Umesh Dayal is a senior researcher at Hewlett-Packard Labs, Palo Alto, California. His current research interests are in distributed information systems, workflow management, data mining, and information management issues related to the emerging global information infrastructure. He received his Ph.D. and S.M. degrees from Harvard University, his M.E. and B.E. degrees from the Indian Institute of Science, and his B.S. degree from Osmania University, India.

Visual Techniques for Exploring Databases

Daniel Keim, University of Munich

For data exploration to be effective, it is important to include the human in the exploration process and combine the flexibility, creativity, and general knowledge of the human with the enormous storage capacity and the computational power of today's computers. Visual database exploration aims at integrating the human in the exploration process, applying its perceptual abilities to the large data sets available in today's computer systems. The basic idea of visual data exploration is to present the data in some visual form, allowing the human to get insight into the data and draw conclusions. Visual data exploration techniques have proven to be of high value in exploratory data analysis and they also have a high potential for exploring large databases. Visual database exploration is especially powerful for the first steps of the data mining process, namely understanding the data and generating hypotheses about the data, but it may also significantly contribute to the actual knowledge discovery by guiding the search using visual feedback.

The goal of this tutorial is to show the potential of visualization technology for exploring large databases. The tutorial provides an overview of the state-of-the-art in data visualization and provides a classification of the existing data visualization techniques. Besides describing each of the classes, the tutorial focuses on new developments in data visualization, which are relevant to the area of knowledge discovery, and describes a wide range of recently developed techniques for visualizing large amounts of arbitrary multi-attribute data which does not have any two- or three-dimensional semantics and therefore does not lend itself to an easy display. A detailed comparison shows the strength and weaknesses of the existing techniques and reveals potentials for further improvements. Several examples demonstrate the benefits of visualization techniques for exploring databases. The tutorial concludes with an overview of existing database exploration and visualization systems, including research prototypes as well as commercial products.

Daniel Keim is one of the leading experts in the field of visual database exploration, and he was the chief engineer in

designing the VisDB system—a visual database exploration system. Keim received his diploma (equivalent to an MS degree) in computer science from the University of Dortmund in 1990 and his Ph.D. in computer science from the University of Munich in 1994. Currently, he is a teaching and research assistant (approximately equivalent to an assistant professor) at the Institute for Computer Science of the University of Munich, Germany.

Tutorial 7: 4:00–6:00 PM

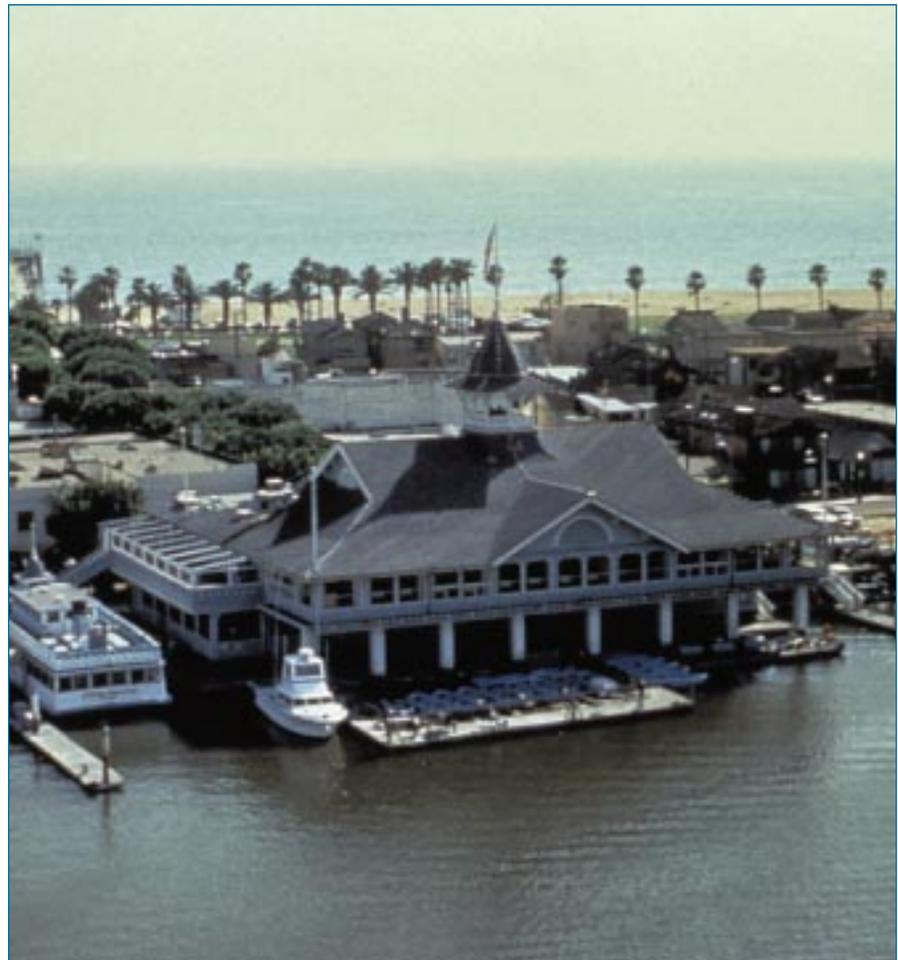
Statistical Models for Categorical Response Data

William DuMouchel, AT&T Research

This tutorial will survey the most common models and methods statisticians use to fit and test relationships among categorical (discrete) data. Most of these techniques are described in statistics texts such as *Categorical Data Analysis*, by Alan Agresti, (Wiley 1990) and are widely available in popular computer packages such as SAS and Splus. Therefore it is al-

most *de rigueur* for someone with a new classification technique to compare the proposal to one or more of these standard methods. The tutorial will focus on log-linear and logistic regression models, and related models such as probit, poisson regression, and survival models. In the short time available, priority will be given to explaining why these techniques are so popular among statisticians, and to how the basic models have been extended to handle variables having more than two categories or when some of the variables have continuous or ordinal scales. Examples of model fitting, model search and model comparison using SAS and Splus will be presented and discussed.

William DuMouchel has been on the faculties of the University of California, Berkeley; University of Michigan; University of London; MIT; and Columbia University. From 1987 to 1992 he was chief statistical scientist at BBN Software Products, helping to design and develop commercial software advisory systems for data analysis and experimental design. He is currently at AT&T Labs – Research, Florham Park, New Jersey.



KDD Workshop—Issues in the Integration of Data Mining and Data Visualization

Sunday, August 17, 8:30 AM – 5:00 PM

Data visualization deals with the effective portrayal of data with a goal towards insight about the data. Typically, the data is of high volume, multidimensional in nature, and does not lend itself to easy display. The data is also often non-spatial and temporal in nature.

Data visualization software systems are very popular with end-user domain scientists who require visual tools to explore and analyze their data. These visual tools however are used strictly as output of the exploration process and have received much attention whereas the input issues to the exploration process still have not. The KDD community is concerned with two aspects of visualization techniques: 1) Its use at the “back-end” of the exploration process to help understand models extracted by data mining algorithms, and 2) Scalability issues in visualization: how do we make it efficient in presence context of large databases where data access is expensive. The visualization community looks at KDD and analytic methods also as applications to generate displays. However, visualization can be used as input to KDD and analytic tools; it can also be used to support computational steering. An effective visualization front-end can guide a data mining algorithm in its search and may result in much better and more easily acceptable solutions. This workshop will continue the discussions started at the first two workshops and focus on these and other issues that make a case for integrating KDD and visualization technologies.

Two previous workshops (Siggraph '90 and Visualization '91) have dealt with areas such as high-level requirements for data structures and access software, and data visualization environments. The first and second workshops on database issues for data visualization were held in 1993 and 1995 and explored the fundamental issues. A number of experimental, prototype, and research systems were presented. The second workshop also saw a beginning interest with data mining and visualization integration. This trend, so significant in the commercial sector today, is in its infancy and is in need of much research attention.

Position statements and papers are welcome on the following issues as they re-

late to KDD and data visualization integration. We would like to keep discussions focused on the end result, which is improving the integration of data mining and knowledge discovery systems with visualization:

- Requirements visualization places on knowledge discovery systems
- Data models and access structures
- Modeling the user—tasks, processes, support issues
- Advanced user interfaces for data mining
- Visual languages for data mining
- System integration issues
- Computational steering for data mining
- Scalability to large databases
- Distributed, heterogeneous data set issues—data and computation sharing
- Examples of integrated systems
- Applications of integrated systems

Additional information about this workshop can be found at the following: www.cs.uml.edu/~grinstei/kddvis-workshop.html

Paper Submissions

(Deadline June 15.) Papers (and position papers to be expanded for final publication) are solicited that present research results in the integration of data mining and visualization. Papers should be limited to 5,000 words and may be accompanied by NTSC video. These should describe some original research on the particular subject, and how it fits in with the overall theme of the workshop. Proper references should be cited.

Registration Fee

Registration forms will be sent to the accepted participants. There is a single registration fee of US \$100, which covers the workshop sessions, preprints, and coffee breaks.

Workshop Organizers

Georges Grinstein
Institute for Visualization and
Perception Research
University of Massachusetts at Lowell
Lowell, MA 01854
E-mail: grinstein@cs.uml.edu
Fax: (508) 934-3551
Phone: (508) 934-3627

Andreas Wierse
Institute for Computer Applications
Department of Computer Simulation
and Visualization
Pfaffenwaldring 27
D-70550 Stuttgart, Germany
E-mail: wierse@rus.uni-stuttgart.de,
Fax: +49(0)711-682357
Phone: +49-711-685-5796

Usama Fayyad
Microsoft Research
Redmond, WA 98052-6399
E-mail: fayyad@microsoft.com
Fax: (206) 936-7329
Phone (206) 703-1528

Demonstrations and Exhibits of Knowledge Discovery Products

Following the success of the demonstration sessions in previous KDD conferences, the KDD-97 program will also include demonstrations of knowledge discovery products, knowledge discovery applications and research prototypes. Unlike previous demonstration sessions, we will clearly differentiate between commercial product demonstrations and research demonstrations.

Exhibits

We invite commercial vendors to exhibit at KDD-97. The exhibitor fee for KDD-97 will be a nominal \$250.00. Exhibitors will be provided with a 6 foot table top. In this space vendors will be allowed to distribute product or company literature, show product demonstrations and set up signage. Vendors will need to bring all necessary hardware and software that they will require for their demonstrations.

The exhibit area will be open August 15th from 12:30–5:00 PM.

Total attendance at KDD-96 was 457. Of these 35 percent were affiliated with universities and 65 percent were affiliated with industry. If you would like to exhibit at KDD-97 please fill out the registration form and send it along with the name of your product(s) and/or service(s) and a 200 word (maximum) description of product(s)/service(s) to:

AAAI
KDD-97 Exhibit
445 Burgess Drive
Menlo Park, CA 94025 USA

Your description will be published in the conference program.

Research Prototype Demonstrations

We are also soliciting demonstrations of research prototypes at KDD-97. This demonstration session will be held on August 15 from 12:30 to 5:00 PM. We have a limited budget for providing hardware for research demonstrations. This year we will give priority to demonstrations that are in conjunction with accepted papers at KDD-97. Within budget and space constraints we will make every effort to accommodate as many demonstrations as possible. If you would like your demonstration to be considered for KDD-97 please provide the following information to Tej Anand (tej.anand@atlanta.ga.ncr.com) by June 1, 1997:

- Name of demonstration
- Title of paper (if this demonstration is in conjunction with a paper/poster at KDD-97)
- Development team
- Affiliations of development team members
- Contact telephone number
- Description of demonstration (approximately 200 words)
- What is unique about your system or application? (No more than 50 words)
- Status: Is the system a research prototype, a commercially available product, or a fielded application?
- Hardware required: Are there any special memory or disk requirements?
- Operating system (specific version number)
- WAN connection needed? (Are there any special modem requirements?)
- Will you bring your own hardware?
- Any other requirements?

Registration Fees

The KDD-97 program registration includes admission to four tutorials, four tutorial syllabi, technical and demo sessions, the opening reception, the *KDD-97 Proceedings* and mid-morning and afternoon coffee breaks. Onsite registration will be located in the foyer outside the California Ballroom, Newport Beach Marriott Hotel and Tennis Club, lobby level.

Early Registration

(Postmarked by June 10)

AAAI Members

Regular \$295 Students \$95

Nonmembers

Regular \$375 Students \$155

Late Registration

(Postmarked by July 15)

AAAI Members

Regular \$350 Students \$125

Nonmembers

Regular \$425 Students \$180

On-Site Registration

(Postmarked after July 15 or onsite.)

AAAI Members

Regular \$400 Students \$150

Nonmembers

Regular \$475 Students \$210

Workshop Registration

Registration forms will be sent to the accepted participants. There is a separate registration fee of US \$100 which covers the workshop sessions, preprints, and coffee breaks.

Payment Information

Prepayment of registration fees is required. Checks, international money orders, bank transfers and travelers' checks must be in US dollars. American Express, MasterCard, VISA, and government purchase orders are also accepted. Registration applications postmarked after the early registration deadline will be subject to the late registration fees. Registration applications postmarked after the late registration deadline will be subject to on-site registration fees. Student registrations must be accompanied by proof of full-time student status.

Refund Requests

The deadline for refund requests is July 25, 1997. All refund requests must be made in writing. A \$75.00 processing fee will be assessed for all refunds.

Registration Hours

Registration hours will be Thursday-Saturday, August 14-16, 7:30 AM-6:00 PM and Sunday, August 17, 8:00 AM-3:00 PM. All attendees must pick up their registration packets for admittance to programs.

Housing

AAAI has reserved a block of rooms at our headquarters hotel—the Newport Beach Marriott Hotel—at reduced conference rates. Conference attendees must contact the hotel directly and identify themselves as KDD-97 registrants to qualify for the reduced rates. Rooms will be assigned on a first-come, first-served basis. All rooms are subject to a 10% occupancy tax.

Newport Beach Marriott Hotel

900 Newport Center Drive

Newport Beach, CA 92660

Phone: (714) 640-4000

Fax: (714) 640-4918

Single: \$105.00 (1 person, 1 bed)

Double: \$115.00 (2 persons, 2 beds)

Check-in time: 4:00 PM

Check-out time: 12:00 noon

Cut-off date for reservations: July 24, 1997.

All reservation requests for arrival after 6:00 PM must be accompanied by a first night room deposit, or guaranteed with a major credit card. The Newport Beach Marriott Hotel will not hold any reservations after 6:00 PM unless guaranteed by one of the above methods. Reservations received after the cut-off time will be accepted on a space or rate available basis.

Reservations accepted without a credit card guarantee or advance deposit are subject to cancellation at 6:00 PM on the day of arrival.

Transportation

Air Transportation and Car Rental

Newport Beach, California – Get there for less! Discounted fares have been negotiated for this event. Call Conventions in America at (800) 929-4242 and ask for Group #428. You will receive five to ten percent off the lowest applicable fares on American Airlines, or the guaranteed lowest available fare on any carrier. Travel between August 11-21, 1997. All attendees booking through CIA will receive free flight insurance and be entered in their bi-monthly drawing for worldwide travel for two on American Airlines! Hertz Rent A Car is also offering special low conference rates, with unlimited free mileage.

Call Conventions in America:

(800) 929-4242, ask for Group #428.

Reservation hours:

M-F 6:30 AM-5:00 PM Pacific Time.

AAAI Press Books on KDD

Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95)

Edited by Usama M. Fayyad and Ramasamy Uthurusamy

Papers in this proceedings focus on unifying themes such as the use of domain knowledge, managing uncertainty, interactive (human-oriented) presentation, and applications.

348 pp.
\$50.00 paperback
ISBN 0-929280-82-2

Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)

Edited by Evangelos Simoudis, Jia Wei Han, and Usama Fayyad

400 pp.
\$55.00 paperback
ISBN 1-57735-004-9

Advances in Knowledge Discovery and Data Mining

Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy

Published by the AAAI Press / The MIT Press

625 pp.
\$50.00
ISBN 0-262-56097-6

To order, consult your registration form, see AAAI Press's web page (www.aaai.org/Press/), or call us at (415) 328-3123

Outside US and Canada:

(619) 453-3686

(619) 453-7679 (fax).

Internet: scitravel@aol.com

24-hour emergency service:

(800) 748-5520.

If you call direct: American (800) 433-1790, ask for index #S 9485.

Hertz: (800) 654-2240, ask for offer CV#24250.

Ground Transportation

The following information provided is the best available at press time. Please confirm fares when making reservations.

Airport Connections

The Newport Beach Marriott Hotel provides complimentary airport transportation to/from John Wayne /Orange County Airport.

Super Shuttle: (714) 517-6600. The one-way fare from the Los Angeles International Airport (LAX) to the Newport Beach Marriott Hotel is \$21.00 per person. Reservations 24 hours in advance are recommended. Discover Card, traveller's checks and cash are accepted.

Taxi

Taxis are available at John Wayne Airport. Approximate fare from the airport to downtown Newport Beach is \$14.00. Orange County Yellow Cab Service: (714) 546-1311. The approximate taxi fare from the Los Angeles International Airport to the Newport Beach Marriott Hotel is \$75.00 – \$80.00.

Bus

Greyhound/Trailways Lines: The depot is located at 100 W. Winston Road, Anaheim, CA 92805. For information on fares and scheduling, call (714) 999-1256.

Rail

The Amtrak (Southern Pacific Railroad) stations are located at Santa Ana, Irvine and Anaheim. For general information and ticketing, call (800) 872-7245.

City Transit System

The Orange County Transit District (OCTD) serves Newport Beach, Balboa Island and Corona del Mar. Basic local fare is \$1.00. For general information call (714) 636-RIDE.

Parking

Parking is available at the Newport Beach Marriott Hotel. The daily rate for valet parking is \$6.00, and \$8.00 overnight. Self-parking is complimentary.





Newport Beach, California!

Newport Beach is located along the beautiful Pacific Ocean in Orange County, California, nestled south of Los Angeles, north of San Diego, southwest of Disneyland in Anaheim, and adjacent to John Wayne/Orange County Airport. Surrounded by one of the largest small-boat harbors in the world and lazily stretching itself along more than six miles of scenic Pacific coastline, Newport Beach beckons national and international visitors to moor at the magnificent harbor and discover “The Colorful Coast.”

Disclaimer

In offering American Airlines, Hertz Rent A Car, Newport Beach Marriott Hotel, and all other service providers, (hereinafter referred to as “Supplier(s)”) for the Third International Conference on Knowledge Discovery and Data Mining, AAAI acts only in the capacity of agent for the Suppliers which are the providers of the service. Because AAAI has no control over the personnel, equipment or operations of providers of accommodations or other services included as part of the KDD-97 program, AAAI assumes no responsibility for and will not be liable for any personal delay, inconveniences or other damage suffered by conference participants which may arise by reason of (1) any wrongful or negligent acts or omissions on the part of any Supplier or its employees, (2) any defect in or failure of any vehicle, equipment or instrumentality owned, operated or otherwise used by any Supplier, or (3) any wrongful or negligent acts or omissions on the part of any other party not under the control, direct or otherwise, of AAAI.

Newport Beach Visitor Information

A concierge desk is available in the Newport Beach Marriott Hotel. They can assist with dining reservations, directions, tour bookings, entertainment suggestions, and transportation information.

Maps and brochures are also available.

Information about Newport Beach is also available on the world wide web:

www.newport.lib.ca.us/NBCVB/NBCVB.htm

(Photographs in this brochure are courtesy, Newport Beach Convention and Visitors Bureau.)

