Data mining is an important tool in science, engineering, industrial processes, healthcare, business, and medicine. The datasets in these fields are large, complex, and often noisy. Extracting knowledge requires the use of sophisticated, high-performance and principled analysis techniques and algorithms, based on sound theoretical and statistical foundations. These techniques in turn require powerful visualization technologies; implementations that must be carefully tuned for performance; software systems that are usable by scientists, engineers, and physicians as well as researchers; and infrastructures that support them.

This conference provides a venue for researchers who are addressing these problems to present their work in a peer-reviewed forum. It also provides an ideal setting for graduate students and others new to the field to learn about cutting-edge research by hearing outstanding invited speakers and attending tutorials (included with conference registration). A set of focused workshops are also held on the last day of the conference. The proceedings of the conference are published in archival form, and are also made available on the SIAM web site.

**TOPICS OF INTEREST**

**Methods and Algorithms**
- Classification
- Clustering
- Frequent Pattern Mining
- Probabilistic and Statistical Methods
- Spatial and Temporal Mining
- Data Stream Mining
- Abnormality and Outlier Detection
- Feature Selection / Feature Extraction
- Dimension Reduction
- Data Reduction
- Mining with Constraints
- Data Cleaning and Noise Reduction
- Computational Learning Theory
- Multi-Task Learning
- Adaptive Algorithms
- Scalable and High-Performance Mining
- Mining Graphs
- Mining Semistructured Data
- Mining Complex Datasets
- Mining on Emerging Architectures
- Text and Web Mining

**Applications**
- Astronomy & Astrophysics
- High Energy Physics
- Collaborative Filtering
- Earth Science
- Risk Management
- Supply Chain Management
- Customer Relationship Management
- Finance
- Genomics and Bioinformatics
- Drug Discovery
- Healthcare Management
- Automation & Process Control
- Logistics Management
- Intrusion and Fraud detection
- Intelligence Analysis
- Biosurveillance
- Sensor Network Applications
- Social Network Analysis
- Application Case Studies
- Other Novel Applications

**Human Factors and Social Issues**
- Ethics of Data Mining
- Intellectual Ownership
- Privacy Models
- Privacy Preservation Techniques
- Risk Analysis
- User Interfaces
- Interestengness and Relevance
- Data and Result Visualization

**WORKSHOPS AND TUTORIALS**
The conference will feature workshops and tutorials on several special topics. Proposals for workshops and tutorials are due on October 12, 2007. Please see the conference web site for submission requirements.

**STEERING COMMITTEE CO-CHAIRS**
Chandrika Kamath, Lawrence Livermore National Laboratory
Vipin Kumar, University of Minnesota

**CONFERENCE CO-CHAIRS**
Chid Apte, IBM Research
Haesun Park, Georgia Institute of Technology

**PROGRAM CO-CHAIRS**
Mohammed Zaki, Rensselaer Polytechnic Institute
Ke Wang, Simon Fraser University

**TUTORIAL CHAIR**
Ian Davidson, SUNY, Albany

**WORKSHOP CHAIR**
Huan Liu, Arizona State University

**PUBLICITY CHAIR**
Pang-Ning Tan, Michigan State University

**IMPORTANT DATES**
Abstract Due: **October 5, 2007**
Manuscripts Due: **October 12, 2007**
Author Notification: **December 21, 2007**
Camera Ready Papers: **January 25, 2008**