

SECOND ANNOUNCEMENT and REGISTRATION FORM

40th Symposium on the Interface:
Computing Science and Statistics

THEME:
RISK : REALITY

21-24 May 2008
Durham Marriott
Durham, NC

<http://www.niss.org/interface2008>

<http://www.interfacesymposia.org>

KEYNOTE SPEAKER:

Siddhartha Dalal
Rand Corporation

"Risk and Public Policy"

INTERFACE 2008

Sponsored by:
Interface Foundation of North America, Inc.

Hosted by:
National Institute of Statistical Sciences

Cooperating Organizations:
ASA, CSNA, ENAR, IASC, IMS, INFORMS, SIAM, WNAR

This symposium is a long-standing forum focusing on the interface between computing science and statistics.

Invitation

The Interface Foundation of North America cordially invites you to participate in the 40th Interface Symposium, the premier annual conference on the interface of computing science and statistics. The Foundation is a non-profit educational corporation founded in 1987 to sponsor the symposium and to publish the proceedings. For further information about IFNA, visit our website at: <http://www.interfacesymposia.org>

The theme of Interface 2008 is RISK : REALITY from *Information to Inference* and from *Technical Developments to Specific Contexts*.

Through the dual focus on the problems of information extraction, risk modeling, analysis and decision-making and on the computational technology and advances in tools to make characterization, quantization and evaluation of risk possible, this symposium will address issues central to understanding real risks and to conceptualizing potential risks and potential risks and risk management scenarios.

The Keynote Speaker is Siddhartha Dalal, Senior Advisor to the President for Technology at Rand Corporation. He brings an extraordinary background of theory, application and vision for statistics, statistical computing and information technology to hard problems modeling reality, developing technology for prediction and setting policy.

Contact Information

Conference Chairs:

Dr. Alan Karr, Director, NISS karr@niss.org
Dr. Nell Sedransk, Associate Director, NISS sedransk@niss.org

Mail: Interface Foundation of North America, Inc.
P.O. Box 7460
Fairfax Station, VA 22039-7460

Phone: (919) 685-9300 Fax: (919) 685-9310

Call for Participation

Invited sessions have been organized by the Program Committee and were designed to cover broadly research related to the theme. See the section below on the Invited Program.

Contributed papers are specifically invited that pertain to any of the bulleted topics listed in the program description below.

Papers are also highly encouraged on advances in computational statistics / computation sciences / statistical computation / statistical graphics; these advances *need not* have particular relevance to RISK.

Abstracts for all sessions should be submitted by 15 April 2008 via the conference website. Papers received by 30 June 2008 will be included in the Interface 2008 Proceedings. Full instructions and other information can be found at the conference website: <http://www.niss.org/interface2008>

Registration and Financial Support

Interface Registration (including the Interface banquet) is \$350 for regular Interface members; \$465 for non-members; and \$150 for students. Members of cooperating societies can register at a discount: \$400. Single-day registration (not including the banquet) is \$175. Guest banquet tickets may be purchased for \$65 each.

Early registration rates (at a discount) apply through 30 March 2008; see registration form.

In order to enroll in a short course, full Interface registration is required. Short course registration is \$165 for members for a single course and \$250 for both (lunch included); student rates are \$75 for one and \$125 for both. See registration form for complete fee schedule. Space is limited and will be filled on a first-come basis.

Registration for either related conferences on 21 May 2008 must be made separately, information is available at: www.niss.org/sdns2008 and at www.samsi.info.

Hotel Accommodations

The Interface 2008 conference hotel is the Durham Marriott at the Civic Center located at 201 Foster Street in Durham, North Carolina 27701. Reservations for Interface 2008 and the related conferences can be made by telephone to: (800) 909 – 8375. The unique conference code for the block of rooms reserved at the special rate of \$109 per night (for all three conferences) is: **NISS**. Please use this code when making your reservation; this special rate will be available until the closing date: 21 April 2008

Transportation

The Raleigh-Durham Airport (RDU) serves the Triangle Area with domestic and international flights. Ground service to Durham (16 miles) is available via taxi or direct shuttle service to the Durham Marriott.

North Carolina Triangle

The Triangle area includes the old communities of Durham, Chapel Hill and Raleigh with their many universities and colleges, particularly Duke University, the University of North Carolina and North Carolina State University respectively. Situated in their midst is Research Triangle Park established as a joint venture and now home to the National Institute of Statistical Sciences and the Statistical and Applied Mathematical Sciences Institute.

Short Courses

Wednesday 21 May 2008

Short Course I: "Omics"

Stan Young, NISS

Kejun (Jack) Liu, OmicSoft

Time 8:00 am – 12 noon

High-dimensional data such as microarray, proteomics and metabolomics data typically have many more variables than samples. Massive genomic data sets, for example, are now coming on line. Such a data set might include hundreds of thousands of patients and on the order of 1 million predictors. Technical problems involve computer science, visualization and statistics issues. This tutorial will cover new matrix factorization methods. We will also explore examples from the microarray area using data drawn from the publicly available NCBI GEO database. There will be hands on use of free demonstration-version software.

S. Stanley Young and Kejun (Jack) Liu

Dr. Young's interests include analysis of complex data sets, algorithms, statistical strategies, and confidence that claims are valid. He has experience in all phases of drug discovery. He is a Fellow of the ASA and an adjunct professor at three research universities.

Dr. Liu graduated from NCSU in statistics and bioinformatics. He has 16 years experience in writing complex analysis software. He is the CTO of Omicsoft, a spin out company from GlaxoSmithKline.

Short Course II: Data Confidentiality

Alan Karr, NISS

Time: 1.30 pm – 5:30 pm

Organizations in every sector--government, industry and academia--face the dilemma of being required, or desiring, to share confidential data. Today's electronic world has exacerbated threats to confidentiality, but also promises new solutions. Multiple disciplines, including statistics, computer science and domain science, are needed to create the solutions. This short course will treat data confidentiality from multiple perspectives: the nature of the problems, "traditional" solutions such as masking data prior to release, and modern and emerging solutions such as remote analysis servers and verification servers. The framing concept is that data confidentiality is a decision problem entailing informed tradeoffs between quantified measures of disclosure risk and data utility.

Alan F. Karr

Alan F. Karr, Director of NISS, has been a leader in data integrity (quality, integration, confidentiality) research and methodology for the past decade. In addition to holding academic credentials as professor at Johns Hopkins, and now at UNC, he is a Fellow of both ASA and IMS and a member of the Board of Governors of Interface Foundation.

Interface 2008 Symposium Theme

RISK : REALITY

Everywhere today RISK is assessed, projected, analyzed, managed. Decisions based on perceived RISK are made daily in every economic sector and, indeed each aspect of life. The pharmaceutical industry continually deals with risk of rare and unexpected side effects of marketed drugs. Homeowners and federal agencies like NOAA deal with the risk of Categories IV and V hurricanes. Security agencies deal with risk of terrorists crossing borders into the US. Federal agencies and companies risk unauthorized penetration of classified and/or proprietary databases. Nuclear power plants must manage risk of radiation escape; while other energy resource managers evaluate strategies to manage risk to the power grid. The focus of Interface 2008 will encompass all these aspects.

Understanding RISK depends crucially on *Information* –

- First - *Acquiring Information*:
 - data acquisition
 - data bases
 - expert opinion
- Second - *Extracting Knowledge* using computational/statistical tools for:
 - massive data / data bases
 - high-dimensional data
 - streaming data
 - text
- Third - *Drawing Inference*:
 - modeling complex systems
 - dynamic analytic processes
 - verification and validation of models and algorithms
 - decision theory

Managing RISK depends on understanding specific *Contexts* –

- extreme events
- natural disasters
- infrastructure (e.g., energy)
- communications & internet
- portfolios
- adversaries & human threats

Program Committee

David Banks (Duke), Dipak Dey (U Connecticut), David Dunson (NIEHS), Lutz Edler (German Cancer Research Center, Heidelberg), Alan Gelfand (Duke), Karen Kafadar (U Indiana), David Marchette (NSWC), Steve Marron (UNC), Wendy Martinez (ONR), George Michailidis(Michigan), Amy Nail (NCSU), Michael Porter (NCSU), David Rios Insua (U Rey Juan Carlos), Yasmin Said (George Mason), Richard Smith (UNC), Michael G. Schimek (U Graz), Jeffrey Solka (NSWCDD), David van Dyk (UC Irvine), Ed Wegman (George Mason), Stanley Young (NISS), Helen Zhang (NCSU)

Local Host

National Institute of Statistical Sciences
Alan F. Karr, Director and Nell Sedransk, Associate Director
(919) 685-9300

Interface 2008

RISK : Reality

Keynote Address

Risk Analysis and Detection of Nuclear Material in Containers Entering US: It's Not Computing Prowess Alone

Siddhartha R. Dalal
Senior Technology Adviser to the President
RAND Corporation

Given the mandate by US Congress that all goods entering US should be inspected for illicit nuclear material, DHS, Department of Homeland Safety, is moving towards one hundred percent inspection of all containers entering US at various ports of entry for nuclear material. Around ninety-five percent of the containers entering at these ports of entry are currently being inspected. This has enabled collection of terabytes of data on millions of containers and their contents. The data include text, radiation portal and other data. Different sets of methods have emerged to analyze each set of data. But, up to now no effort has been made to analyze these sets of data in a unified manner for detecting illicit nuclear material. In this talk we discuss a number of challenges in creating a real time unified decision system from a risk analysis perspectives and the corresponding challenges in data analysis. The methodologies proposed here are based on a number of machine learning and statistical approaches and are generalizable to all situations where screening is involved.

Siddhartha Dalal

Siddhartha Dalal is the Senior Technology Adviser to the President at RAND. Sid's industrial research career began at Math Research Center at Bell Labs followed by Bellcore/Telcordia Technologies. Most recently he was a vice president of research at Xerox. He has co-authored over seventy publications, several patents and two NRC reports covering the areas of risk analysis, econometrics modeling, image processing, stochastic optimization, data/document mining, software engineering and Bayesian methods. He is recipient of several best paper awards from IEEE, ASA and ASQC.

Invited Program

Probabilistic Models in Risk Assessment

Organizer: David Banks

Mehmet Sahinoglu	Troy U	<i>Security Risk for Computer Systems</i>
Alyson Wilson	LANL	<i>Bayesian Reliability Analysis</i>
David Banks	Duke U	<i>Adversarial Risk Analysis</i>

Modeling of Extreme Events and Analysis of Risk

Organizers: Dipak Dey & David Rios-Insua

Elijah Gaioni	U Conn	<i>Semiparametric Functional Estimation using Quantile-based Prior Elicitation</i>
Sourish Das		<i>Hurricane Activity in Context of Changing Environment</i>
Jesus Rios	SAMSI & Aalborg U	<i>Risk Analysis for Auctions</i>

Invited Program

Multivariate Extremes

Organizer: Richard Smith

TBA

Spatial Risk Mapping: Prediction and Change Detection

Organizer: Michael Porter

Jason Dalton	SPADAC	<i>Space-time Forecasting of Extreme Events in Complex Environments</i>
Ronald D. Fricker, Jr.	Naval Postgraduate School	<i>Using the Repeated Two-sample Rank Procedure for Detecting Anomalies in Space and Time</i>
Michael Porter	SAMSI & NCSU	<i>A Martingale Methodology for the Quick Identification of Point Process Anomalies</i>

Air Pollution Risk Assessment: from Research to Regulation

Organizer: Amy Nail

Allen Lefohn	ASL & Associates	<i>Realistic Biological and Exposure/Dose Relationships: How They Modify Perceived Human Health & Ecological Risk</i>
Roger Peng	Johns Hopkins U	<i>Statistical Methods for Assessing the Health Risks of Particulate Matter Components</i>
Yongku Kim	SAMSI	<i>How Changing the Ozone Standard Might Affect Respiratory Mortality</i>

Statistical and Computational Issues in Analyzing Sensor Networks

Organizer: Alan Gelfand

George Michailidis	U Michigan	<i>Robust Target Detection & Localization in Wireless Sensor Networks</i>
Carol Y. Lin	CDC	<i>Statistical Issues in Designing an Optimal Detection System with Multiple Heterogeneous Sensors</i>
Soumendra Lahiri	Texas A&M	<i>Analysis of Microsensor Networks from a Statistical Perspective</i>

Sensor Networks and Statistics - New Researchers Session

Organizer: George Michailidis

Sheela Nair	UCLA	<i>Fault Detection for Embedded Networked Sensing</i>
Natalia Katneka	U Michigan	<i>A Cost-efficient Approach to Wireless Sensor Network Design</i>
Gavino Puggioni	Duke U	<i>Analyzing Space-time Sensor Network Data under Suppression and Failure in Transmission</i>

Enhancing Knowledge and Assessing Risk through Analysis of Massive Data

Organizer: Karen Kafadar

Ginger Davis	UVA	<i>Statistical Methods for Detecting Computer Attacks from Streaming Internet Data</i>
Amy Braverman	JPL/CalTech	<i>Massive Data Set Analysis for NASA's Atmospheric Infrared Sounder</i>
Michael Trosset	Indiana U	<i>What Kind of Knowledge Does Locally Linear Embedding Extract?</i>

Invited Program

Change Detection in Random Graphs

Organizer: David Marchette

David Marchette	NSWC	<i>Detecting Activity Changes in Graphs</i>
Youngser Park	Johns Hopkins U	<i>Scan Statistics in Hypergraphs</i>
Elizabeth Beer	Johns Hopkins U	<i>Torus Graph Inference for Detection of Localized Activity</i>

Statistics and Modern Image Analysis, I

Organizer: Steve Marron

S.M. Pizer	UNC	<i>M-reps, Curved Feature Space, Bayesian Segmentation</i>
R. E. Broadhurst	UNC	<i>Quantile Functions for Texture Analysis and M-rep Segmentation</i>
Suman Sen	UNC	<i>Manifold SVM for M-rep Data</i>

Statistics and Modern Image Analysis, II

Organizer: Steve Marron

Brad Davis	Kitware & UNC	<i>Smoothing over Diffeomorphisms</i>
Hongtu Zhu	UNC	<i>Intrinsic Regression Model for Positive Definite Matrices</i>
Haipeng Shen	UNC	<i>Supervised Singular Value Decomposition for Independent Component Analysis of fMRI</i>

Statistics and Evolutionary Biology, I

Organizer: Haipeng Shen

Joel Kingsolver	UNC	<i>Evolutionary Analyses of Function-valued Traits</i>
Travis Gaydos	UNC	<i>Quantification of Curves' Variation and Simplicity to Find Genetic Constraints</i>
Brian O'Meara	National Evolutionary Synthesis Center	<i>Extending Models of Character Coevolution</i>

Statistics and Evolutionary Biology, II

Organizer: Haipeng Shen

Christina Burch	UNC	<i>Distribution of Mutation Effects and Adaptation in an RNA Virus</i>
Mihee Lee	UNC	<i>Deconvolution and Sieve Estimation of Mutation Effect Distribution</i>
Paul Magwene	Duke U	<i>Estimation of Modularity in Biological Systems</i>

Assessing Health Risk from Complex Data

Organizer: David Dunson

Joseph Ibrahim	UNC	<i>A Bayesian Hidden Markov Model for Motif Discovery through Joint Modeling of Genomic Sequence and ChIP-chip Data</i>
Jason Fine	UNC	<i>Analysis of Left-truncated Semi-competing Risks Data with Application to Disease Registries</i>
Lianming Wang	NIEHS	<i>Semiparametric Bayes Modeling of Onset and Progression from Current Status Data</i>

Invited Program

Model-based Risk Assessment in Life Science

Organizer: Lutz Edler

C. Portier

*Finding the Right Path: Using Structurally-Enhanced Pathway
Enrichment Analysis to Identify Targets for High-Throughput Screening*
Data Gaps and Needs in Model-based Risk Assessment
TBA

Lutz Edler
TBA

SNP Analysis Methods and Software

Organizer: Stan Young

Danyu Lin
Kejun (Jack) Liu
TBA

UNC
OmicSoft

HapStat
SNP Analysis Methods
TBA

Risk of Reaching False Conclusions

Organizer: Stan Young

Robert Obenchain
Patrick Ryan
TBA
Alice White

SoftRx
GlaxoSmithKline
TBA
GlaxoSmithKline

TBA
TBA
Discussant
Discussant

Recent Developments in Machine Learning and Classification - to appear in the *Journal of Computational and Graphical Statistics*

Organizer: David Van Dyk

George Michailidis

U Michigan

*An Iterative Algorithm for Extending Learners
to a Semi-supervised Setting*

Tong tong Wu

U Maryland

An MM Algorithm for Multicategory Vertex Discriminant Analysis

Han-Ming (Hank) Wu

Tamkang U

Kernel Sliced Inverse Regression with Applications to Classification

New Developments in Machine Learning and Statistical Modeling for Massive Data

Organizer: Helen Zhang

Jerry Zhu

U Wisconsin

Online Semi-supervised Learning

Yufeng Liu

UNC

Robust Large-margin Classifiers

Howard Bondell

NCSU

Simultaneous Feature Selection and Structure Identification for ANOVA

Text Data Analysis

Organizer: Jeffrey Solka

Elizabeth Hohman

*Generalization of the Vector Space Model
for a Streaming Corpus of Text Documents*

Randall Giles

VCU

Interactive Text Mining with Iterative Denoising

Avory Bryant

Cross Corpus Discovery via Nearest Neighbor Change-point Analysis

Invited Program

Integration of Disparate Types of Information

Organizer: Wendy Martinez

Carey Priebe	Johns Hopkins U	<i>Disparate Information Fusion: On the Exploitation of Multiple Disparate Dissimilarities</i>
Brent Castle	Indiana U	<i>Combining Disparate Information by Nonmetric Multidimensional Scaling</i>
Jeffrey Solka	NSWCDD	<i>Disparate Information Fusion on Images and Text</i>

TBA

Organizer: Yasmin Said

TBA

TBA

Organizer: Edward Wegman

TBA

TBA

Organizer: Edward Wegman

TBA

Conference Schedule

DATE	EVENTS
<p>Wednesday 21 May</p>	<p style="text-align: center;">Related Conferences</p> <p style="text-align: center;">ASA – QMDNS niss.org/sdns2008 SAMSI – RISK samsi.info</p> <p style="text-align: center;">Interface 2008 RISK : Reality</p> <p style="text-align: center;">Short Course I: “Omics” Stanley Young (NISS) & Kejun (Jack) Liu (OmicSoft) (8:00 am – 12 noon)</p> <p style="text-align: center;">Short Course II: Data Confidentiality: Modern Problems, Modern Tools Alan F. Karr (NISS) (1:30 pm – 5:30 pm)</p> <p style="text-align: center;">Board Dinner and Meeting (6:00 pm)</p> <p style="text-align: center;">Interface 2008 Registration & Evening Mixer (8:00 pm – 10:00 pm)</p>
<p>Thursday 22 May</p>	<p style="text-align: center;">Interface 2008 Registration (7:30 am - 5:30 pm)</p> <p style="text-align: center;">Keynote Address Siddhartha Dalal (Rand Corporation) “Risk and Public Policy” (8:30 am - 10:00 am)</p> <p style="text-align: center;">Technical Sessions (10:30 am - 5:30 pm)</p> <p style="text-align: center;">Conference Banquet (7:00 pm - 10:00 pm)</p>
<p>Friday 23 May</p>	<p style="text-align: center;">Interface 2008</p> <p style="text-align: center;">Technical Sessions (8:00 am - 5:30 pm)</p>
<p>Saturday 24 May</p>	<p style="text-align: center;">Interface 2008</p> <p style="text-align: center;">Technical Sessions (8:00 am – 12 noon)</p>

Interface 2008

Registration Form

40th Meeting of the Interface, INTERFACE 2008, 21-24 May 2008.

Name (as it will appear on your nametag) _____

Affiliation (as it will appear on your nametag) _____

Address (work or home) _____

City _____ **State or Province** _____

Zip or Postal Code _____ **Country** _____

Telephone _____ **email** _____

Join Interface now (attend at Member rate) \$35 membership _____

Conference Fee (including Interface Mixer, 1 Banquet ticket, Interface 2008 Proceedings) _____

	By 30 March 2008	On/after 31 March 2008
Interface member	\$300	\$350
Cooperating Society Member	\$350	\$400
Non-member	\$415	\$465
Student	\$125	\$150

(Students: Attach a letter from department chair or major professor to registration form)

Single Day Fee (without Banquet, without Interface 2008 Proceedings) _____

	By 30 March 2008	On/after 31 March 2008
<input type="checkbox"/> 22 May <input type="checkbox"/> 23 May <input type="checkbox"/> 24 May	\$150	\$175

Short Course(s) "Omics" only Data Confidentiality only Both _____
(registration for at least single day required)

	One Short Course	Two Short Courses (lunch included)
Interface member	\$165	\$250
Cooperating Society Member	\$165	\$250
Non-member	\$225	\$375
Student	\$ 75	\$125

Guest Banquet Tickets By 30 March 2008 On/after 31 March 2008
Number of Tickets _____ @ \$150 \$175 _____

Total _____

Make checks payable to: Interface. Mail to Interface 2008, P.O. Box 7460, Fairfax Station, VA 22039-7460 USA.
By credit card, fax to: (703) 993-1700 (attention: Ms. Elizabeth Quigley).

VISA MasterCard Discover Expiration ___ / ___
Card Number (with 3-digit PIN on reverse of card) _____ / _____

Cardholder's Name _____ Signature _____

Cardholder's Address (if different than above) _____

Email Address for confirmation IF cardholder is not attendee _____

Inquiries to Ms. Elizabeth Quigley at (703) 993-9107 or via email to equigley@gmu.edu.

Returned checks or denied credit cards will be rebilled with \$30 surcharge.