



National Institute of Statistical Sciences
PO Box 14006, Research Triangle Park, NC 27709-4006
Tel: 919.685.9300 FAX: 919.685.9310
www.niss.org

NISS Affiliates
NISS/SAMSI University Affiliates
First Annual Problem Day

Thursday, March 4, 2004

NISS Headquarters, Research Triangle Park, NC

Summary

1 Summary

The first Annual Problem Day of the NISS Affiliates and NISS/SAMSI University Affiliates was held on March 4, 2004 at NISS headquarters in Research Triangle Park, NC, beginning at 9:00 AM.

Present were David Banks (Duke University), Somesh Chattopadhyay (Florida State University), Willa Chen (Texas A&M), Steven Cohen (Bureau of Labor Statistics), Rob Creecy (Census Bureau), Thomas Ferryman (Pacific Northwest National Laboratory), Thomas Gerig (NISS), Alan Karr (NISS), Ravi Khattree (Oakland University), Gary McDonald (NISS), Paul Speckman (University of Missouri), Cliff Spiegelman (Texas A&M), Dongchu Sun (University of Missouri), John Wierman (Johns Hopkins), Michiel van Rhee (ICAGEN) and Stanley Young (NISS).

2 Background

The purpose of the Problem Days is to allow corporate, government and national laboratory affiliates, and NISS itself, to present problems for which they wish to form collaborative relationships, especially with NISS/SAMSI University Affiliates. Such collaborations might lead, for example, to proposals that enable high-impact research to address the problems. The Affiliates Proposal Development Fund (see www.niss.org/affiliates/apdf/apdf.html) is one mechanism by which collaborations can be moved in this direction.

Direct involvement for NISS in Problem Day-generated collaborations is not mandatory, but it is available in cases, where the scale, geographical distribution or other characteristics of the effort (for example, shared postdoctoral appointments) make it sensible. When NISS is the lead organization on a proposal, all subawards carry full indirect costs, so there is no financial penalty for participation in a NISS-led project.

A pre-Problem Day example of this process took place at the 2000 affiliates planning meeting. Jon Kettenring, as affiliates representative from Telcordia, articulated the problem of data quality. This led to continuing research on data quality at NISS that has involved more than ten affiliates, including universities.

3 Presentations

The following presentations were made:

Alan Karr, NISS: Secure Analyses of Distributed Data

Thomas Ferryman, Pacific Northwest National Laboratory (PNNL): BIG Problems from PNNL

David Banks, Duke University and SAMSI: Problems from the SAMSI Program on Data Mining and Machine Learning (DMML)

Alan Menius, GlaxoSmithKline: *omics

Michiel van Rhee, ICAGEN: High-Throughput Screening Data—Pushing Cheminformatics to the Limits

Alan Karr, NISS: Statistical Computing Infrastructure Initiative (SCI²)

Copies of many presentations are available at www.niss.org/affiliates/problemday200403.html.

4 Discussion

A number of additional points and problems were raised during the open discussion that concluded the day:

- How to pursue statistical problems identified in studies conducted by the National Academies. A specific example cited was the recent study *Weighing Bullet Lead Evidence* (see books.nap.edu/catalog/10924.html).
- Principled ways to construct, and deal with the effects of, *ad hoc* changes to surveys.