NISS

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 and NISS/SAMSI University Affiliates PLANNING MEETING SUMMARY Friday, March 7, 2003

1 Summary

The annual Planning Meeting of the NISS Affiliates and NISS/SAMSI University Affiliates was held on March 7, 2003 at NISS headquarters in Research Triangle Park, NC, beginning at 9:00 AM.

Present were David Banks (Food and Drug Administration), James Berger (SAMSI), James Cavendish (General Motors), Lawrence Cox (National Center for Health Statistics), Edward Davis (UNC Chapel Hill—Biostatistics), Linda Davis (American University), Jan de Leeuw (UCLA), John Eltinge (Bureau of Labor Statistics), Valeri Fedorov (GlaxoSmithKline), Thomas Gerig (NISS), Nick Hengartner (Los Alamos National Lab), Alan Karr (NISS), Jon Kettenring (Telcordia Technologies), Ravi Khattree (Oakland University), Ken Kirsten (SPSS), Walter Liggett (NIST), Gary McDonald (NISS), Alan Menius (GlaxoSmithKline), Daniel Naiman (Johns Hopkins), Sastry Pantula (North Carolina State), Avi Singh (RTI International), Keith Soper (Merck & Company), Cliff Spiegelman (Texas A&M), Dalene Stangl (Duke), Dongchu Sun (Missouri), Tommy Wright (Census Bureau), and Stanley Young (NISS).

2 Program Review

Following a welcome and participant introductions, Karr reported on activities, recent initiatives and plans for the Affiliates Program.[1] Major points:

1. Three Assistant Directors of NISS have been appointed, all of whom are involved deeply with the Affiliates Program: Thomas Gerig (NISS/SAMSI University Affiliates), Gary McDonald (Program Development) and Stanley Young (Bioinformatics).

2. SAMSI has led to additional program benefits and new program diversity with the formation of NISS/SAMSI University Affiliates.

[1]The presentation and other meeting-related materials are available at www.niss.org/affiliates/planningmeeting200303.html.

3. Currently, approximately 50 corporations, federal agencies, national laboratories and university departments (listed in Appendix A) are Affiliates.

4. Reflecting the strong influence of the Affiliates Program on NISS, research involving Affiliates currently addresses *data confidentiality* (BLS, BTS, Census, NASS, NCES, LANL, CMU, Purdue, SMU), *data quality* (BTS), *computer model evaluation* (GM, Duke) and *extreme value theory for global climate change and atmospheric pollution* (EPA, Duke, UNC). There is also a Focused Research Group proposal on *computational algebraic statistics* (CMU and other universities).

5. Affiliates events during 2002–03 included:

   - High Throughput Screening Technology Day: October 25, 2002
   - Computer Experiments Short Course: November 7–8, 2002
   - Proteomics Technology Day: March 6, 2003
   - Internet Tomography Technology Day: March 28, 2003
   - EURANDOM Statistical Data Mining Workshop: April 23–24, 2003

   as well as seven SAMSI workshops, this Planning Meeting and the JSM Affiliates Meeting on August 11, 2002.

6. The SAMSI 2003–04 Synthesis Program on Data Mining and Machine Learning is a direct result of support for the program shown at the March 2002 Affiliates Planning Meeting. Affiliates will be invited to provide testbed data sets for the program.

Karr announced creation of the NISS Affiliates, and NISS/SAMSI University Affiliates, Proposal Development Fund (APDF). Two awards are expected annually, of up to $10,000, for expenses associated with preparation of cross-disciplinary, preferably multi-institution proposals by researchers from Affiliates. Requests for APDF support should be for projects that are cross-disciplinary, bold and of sufficient scale to achieve significant impact. Initial proposals are due July 1, 2003, and awards will be announced at the NISS JSM 2003 Reception in San Francisco (on Monday, August 4, 2003). The complete APDF announcement appears in Appendix B.

Gerig, McDonald and Young discussed additional Affiliates initiatives:

- The University Affiliates Campaign;
- Potential involvement of NISS and Affiliates with data to be collected under the TREAD (Tire Recall, Enhancement, Accountability and Documentation) Act;
- A library of Affiliates (Affiliate-provided, with access restricted to Affiliates) data sets.

Discussion occurred throughout the presentations. One recurring theme was the desirability of making NISS and SAMSI events accessible (via the Web or other means) to those not able to attend in person.

3 SAMSI

Berger summarized current and future activities of SAMSI, with particular focus on:

- 2002–03 programs on Inverse Problems, Stochastic Computation and Environmental Modeling.
- 2003–04 programs on Internet Traffic (Fall 2003), Data Mining and Machine Learning (entire year) and Multiscale Modeling and Control Design (Spring, 2004). Additional information, including complete listings of workshop dates, is available at www.samsi.info/200304/programs200304.html.
- Anticipated 2004–05 programs on Social Sciences, Gene Regulation and Data Assimilation.
- The multiple opportunities for Affiliate engagement with SAMSI, which include proposing and leading programs, visits to SAMSI (short- and long-term), workshop participation, providing testbed problems and data sets, postdoctoral fellowships (2+ years) and SAMSI University Fellowships. Graduate students can visit, attend workshops and participate in special activities.
- The 2003 Industrial Mathematical and Statistical Modeling Workshop, to be held July 21-30, 2003 at North Carolina State University. This 10-day workshop for graduate students emphasizes the team approach to multidisciplinary projects and the SAMSI philosophy of merging significant statistical and mathematical components. Additional information is available at www.ncsu.edu/crsc/immw/.

Further information about SAMSI is available on the SAMSI Web site—www.samsi.info, which is also linked from the NISS home page.

4 $2 + \varepsilon$ Minute Madness

Every attendee took (approximately) two minutes to describe concerns, needs and capabilities of his or her organization that are relevant to the Affiliates Program. Corporate, government and national lab affiliates concentrated on personnel and research needs—recurring themes were data quality, data confidentiality, data mining and computer model evaluation, while university affiliates emphasized the research interests of their faculties.

5 Breakout Discussions and Reports

Hour-long breakout discussions were held, addressing questions posed by Karr during his report:

Communication: How can Affiliates provide input (to NISS) most effectively?

Initiatives: What new Affiliates initiatives should be undertaken?

Bridging: How can the program bridge more efficiently between university and corporate/government/national lab Affiliates?

Balance: What is the right balance between activities that serve all Affiliates and those that serve targeted subsets?

The first and fourth of these were addressed by a single group.

Each breakout group reported to the full set of attendees. Key points in the reports and ensuing discussions were:

Communication: Karr reported that items discussed included visits by NISS staff to Affiliates (positive reaction with the group), creation of databases faculty interests and expertise among University Affiliates (positive reaction) and problems raised by any Affiliate (skeptical reaction, given the complexity of many problems), and ways of disseminating events such as Technology Days (ranging in effort and expense from speakers' presentations being placed on the Web to DVDs containing an entire event). Another idea was a consolidated technical report series.

Initiatives: McDonald listed five potential workshop themes identified by the group:

- Articulation of new software capabilities for statistical analysis;
- Critique of the quality of existing statistical computing software (a role that one participant noted is filled in part by the *Journal of Statistical Software*);
- Detection of anomalies in large streams of data;
- Data fusion;
- Detection of bioterrorism.

It was noted that some of these lead to potential new Affiliates.

Bridging: Gerig reported four major recommendations of the group:

1. A "brokering" role for NISS to create initial contacts between faculty from University Affiliates and corporate/government/national lab affiliates;
2. Sharing of postdoctorals, who would be appointed by NISS but spend the majority of their time at an Affiliate;
3. A faculty interest database (see **Communication** above);
4. Improved event dissemination (see §2).

Balance: Karr reported a clear consensus in the group that depth requires that events be targeted. Other items discussed included clearer definition of the vision and purposes of Technology Days by the organizers, tutorials on the afternoon before Technology Days and week-long, very-in-depth, meetings ("summer schools").

Additional items emerging in the discussion were NISS' disseminating information about emerging funding activities and working on behalf of the community to generate such opportunities (in the context, for example, of national defense and homeland security).

The meeting adjourned at 3:15 PM.

A Affiliates as of January, 2003

Corporations: Amgen, Avaya Labs, General Motors, GlaxoSmithKline, Merck & Company, MetaMetrics, Pfizer, Quintiles Transnational, Research Triangle Institute, SAS Institute, SPSS, Telcordia Technologies

Federal Agencies: Bureau of Labor Statistics, Bureau of Transportation Statistics, Census Bureau, Environmental Protection Agency, National Agricultural Statistics Service, National Center for Education Statistics, National Center for Health Statistics, National Institute of Standards and Technology, National Security Agency

National Laboratories: Los Alamos National Laboratory, Pacific Northwest National Laboratory

University Departments: American (Mathematics and Statistics), Carnegie Mellon (Statistics), Duke (Mathematics; Institute of Statistics and Decision Sciences), Emory (Biostatistics), Georgia (Statistics), Iowa (Statistics), Iowa State (Statistics), Johns Hopkins (Mathematical Sciences), Maryland (Mathematics), Maryland Baltimore County (Mathematics and Statistics), Michigan (Statistics + Biostatistics), Minnesota (Biostatistics), Missouri-Columbia (Statistics), North Carolina State (Mathematics; Statistics), North Carolina at Chapel Hill (Biostatistics; Mathematics; Statistics), Oakland (Mathematics and Statistics), Ohio State (Statistics), Pennsylvania State (Statistics), Pennsylvania (Statistics), Purdue (Statistics), Southern Methodist (Statistical Sciences), Stanford (Statistics), Texas A&M (Statistics), UCLA (Statistics + Statistical Consulting Center)

B APDF Announcement

NISS Affiliates
NISS/SAMSI University Affiliates
Proposal Development Fund

March 2003

NISS is pleased to announce creation of the NISS Affiliates, and NISS/SAMSI University Affiliates, Proposal Development Fund (APDF).

Summary. In pursuit of its mission, on behalf of the statistical sciences community, "to catalyze ... high-impact, cross-disciplinary research involving the statistical sciences," NISS expects to make two awards annually, of up to $10,000, for expenses associated with preparation of cross-disciplinary, preferably multi-institution proposals by researchers from NISS Affiliates or NISS/SAMSI University Affiliates. The resulting proposals will be to Federal funding agencies such as the NSF, NIH or DARPA, to private foundations, or to corporations. They will, in general, be submitted by Affiliates, but may if appropriate be submitted by NISS.

APDF awards may be used to support documented expenses, *other than PI and co-PI time*, that further proposal development. Examples of allowable costs include workshops, travel, acquisition of testbed data and graduate student time for pilot investigations.

The APDF seeks projects that are cross-disciplinary, bold and of sufficient scale to achieve significant impact. Selection criteria, described in more detail below, include the cross-disciplinary and collaborative nature of the research, as well as the value added of NISS support. Current NISS projects on data confidentiality (see www.niss.org/dg) and software engineering are examples that fit these criteria.

Eligibility. The PI on the proposal to the NISS APDF must be a statistical scientist associated with a NISS Affiliate or NISS/SAMSI University Affiliate, and must intend to be PI on the resulting proposal. Co-PIs need not be statistical scientists, nor need they be associated with Affiliates.

Matching Requirement. The home institutions of the PI and co-PIs must collectively provide matching at least equal to the funds requested from the APDF. In-kind matching, for example, PI time in the form of course release or summer salary, is allowed and encouraged.

Proposal Details. Proposals to the APDF consist of:

- A five-page description of the proposal to be prepared, including its scientific goals, project personnel and the nature and extent of cross-disciplinary collaboration. The organization (and, if appropriate, program) to which the proposal will be submitted should be identified.

- Two-page, NSF-style biographical sketches of the PI and co-PIs;
- Letter(s) of commitment for the matching requirement;
- Budget.

Proposal Submission. Proposals should be submitted electronically (PDF is the preferred format, but Microsoft Word and PostScript are acceptable.) to apdf@niss.org.

Initial proposals, for awards effective during the academic year 2004–05, are due on July 1, 2003. Awards will be announced at the NISS/SAMSI JSM Reception in August, 2003. In subsequent years, proposals will be due on May 1.

Evaluation Criteria. Proposals will be evaluated by a Committee of the Board of Trustees of NISS, chaired by the Director, on the basis of the following criteria:

- Scientific strength, including the importance of the problems addressed and the potential impact of the research;
- Genuine cross-disciplinarity, including active participation of co-PIs from the applied mathematical sciences (in the spirit of SAMSI) or other scientific disciplines;
- Time commitment of the PI and co-PIs;
- The extent to which APDF support would make something happen that would not happen otherwise.

Preference will be given to proposals that involve young researchers (particularly postdoctorals) and multiple institutions (especially Affiliates), as well as to proposals that intersect NISS' own programs of cross-disciplinary research.