

**Notes from the CANSSI Scientific Planning meeting, 3:30- 4:30 pm, June 13, 2015
Dalhousie University**

Chair: Nancy Reid, University of Toronto (Director, CANSSI)

Present:

John Braun, University of British Columbia, Okanagan (Deputy Director, CANSSI)
Hugh Chipman, Acadia University (CANSSI Associate Director, Atlantic Region)
Mary Thompson, University of Waterloo (CANSSI Associate Director, Ontario)
Alexandre Leblanc, University of Manitoba (CANSSI Associate Director, Manitoba, Saskatchewan, NWT and Nunavut)
Will Welch, University of British Columbia (CANSSI Associate Director, British Columbia, Alberta and Yukon)
Hong Gu, Dalhousie University
Jeffrey Picka, University of New Brunswick
Jean-François Plante, HEC Montréal
Shirley Mills, Carleton University
David Thomson, Queen's University
Jamie Stafford, University of Toronto
Hanna Jankowski, York University
Ejaz Ahmed, Brock University
Angelo Canty, McMaster University
Julie Horrocks, University of Guelph
Changbao Wu, University of Waterloo
Abdulkadir Hussein, University of Windsor
Jim Stallard, University of Calgary
Jason Loeppky, University of British Columbia and UBC Okanagan
Derek Bingham, Simon Fraser University

The following topics had been put forward for discussion at the meeting, and the discussion for each is summarized in italics.

1. Can CANSSI serve a supporting role in undergraduate and graduate training?

- what experience can our Collaborative Research Teams (CRTs) offer about online courses
- could we establish a web resource for data science programs
- what should we be doing in health-related training

Derek Bingham (SFU), lead investigator of a CRT, described an online course put on by the team at UBC, SFU and Acadia universities. The course included guest speakers from outside these universities. The course was judged to be very successful, and it was felt that more inter-university courses would be beneficial. The technology need not be expensive: a WebEx license might be as little as \$50 per month. Slides from an online course can be made available. Nancy Reid noted that it may also be feasible to record and archive the course, as was done with the graduate courses at Fields that were part of the Big Data thematic program in the first half of 2015.

Support was expressed for the idea of a web resource for data science programs, but it was also cautioned that the resource would be effective only if kept up to date. Links to useful courses, talks, papers etc. could be maintained, perhaps by universities in partnership. The webpage of U of Toronto PDF Roger Grosse is an interesting example.

The question was raised of whether data science educational resources are part of CANSSI's mandate, since CANSSI is focused on the advancement of research, and since we need to spend our time strategically. Some points were heard arguing in favour of inclusion:

- it needs to be communicated to young students that statistics is a core aspect of data science*
- CANSSI needs to be concerned with training at university and pre-university to try to help ensure that students receive the right kind of background to do research in data science*
- NSERC funds research at the institutes, but is looking for evidence of active commitment to educational outreach*

2. Can CANSSI be a resource for coordinating postdoctoral fellowship opportunities and applicants?

- if yes, what is needed to get started on this*
- how would it interface with other similar efforts, e.g. SSC jobs, or d-ssc mailing list*
- what would be most helpful for potential supervisors and for potential pdfs*

Mary Thompson noted that CANSSI funds postdoctoral fellowships (PDFs) within the CRTs and also outside them; however, we are currently limited to partial funding for three or four per year. PDFs come to their positions with statistical science expertise at the doctoral level, and also play key roles in supporting collaborations by putting significant time into learning about the collaborating disciplines and helping to make the team cohesive. They provide this kind of value whether at a small or a large university. Thus an additional role for CANSSI would be to help coordinating opportunities. Currently, we do advertise on the CANSSI website notices of PDF positions that come our way. We can keep trying to make the case to granting agencies that they fund more PDF positions. We might be able to facilitate matches from industry and other organizations.

Support was expressed for CANSSI seeking further ways of funding PDFs, and facilitating opportunities. As well:

- it was suggested that we might set up a portal for candidates to indicate their interest in finding a suitable PDF position; Math Jobs often has this function, but it is not clear that statistical science students make extensive use of it*
- support was expressed for CANSSI funding some inter-institutional PDFs outside the CRTs*
- it was noted that in Canada, while a PDF position may have important career benefits, many PDF stipends are not competitive with those of other opportunities; PDF stipends in the US tend to be much higher*

3. How can CANSSI best work with Mitacs? Note that Mitacs has just signed an MOU with NSERC, which will coordinate Mitacs programs with NSERC RPP programs: Engage Grants, CRD, and CREATE

- Many of our CRTs are already taking advantage of the Accelerate internship program; could we leverage our PDF program through the Elevate Industrial Postdoc program?; and should we encourage CANSSI PDFs to take part in Step workshops (which build soft skills)?
- CANSSI and Mitacs could establish a small working group to develop a strategy for engaging researchers in business analytics, biotechnology, and risk management with those deeply involved in computational advances, visualization tools, and design and analysis of data. We can certainly see the scope for joint postdoctoral fellowships in financial services, insurance, graphics, and visualization, bioinformatics, and climate sciences, for example.
- expanding training around big data and data science.
- many graduate programs in data science will involve an internship component, which could (should?) be routed through Mitacs.
- a version of an industrial problem-solving workshop activity might be developed jointly with Mitacs in the form of “data-thons”, modelled after the hackathons very popularly used in computer science and in finance.

In the discussion it was heard that:

- *now for certain types of project not-for-profit organizations are eligible as Elevate (two-year PDF) or Accelerate (internship) partners*
- *the CANSSI website might include a directory for MITACS projects – those already in place and those seeking trainees*
- *there could be a portal for candidates seeking internships and PDFs*
- *the CANSSI newsletter might point to MITACS opportunities and feature success stories*

4. Science: we now have 6 CRTs, and another competition is underway. Next year we expect to have 8, of which two will be in their last year; in steady-state we expect to have 6 CRTs. Topics of the current CRTs include: copula modeling, state-space models and fisheries science, computer modelling in physical sciences, point process modelling and wildfire regimes, complex survey data, spectral methods in physical sciences. Should we encourage/steer research in other areas? What are the big gaps?

This topic was not discussed, the hour having been taken up with the first three.