



CANSSI's Strategic Plan – Draft Version

Vision Statement

Canada's catalyst for discovery and innovation in the statistical and data sciences.

Mission Statement

CANSSI is the national institute that advances the development, application, and communication of cutting-edge statistical and data science research and training. CANSSI builds multi-disciplinary collaborations with leadership in statistical science and data science. CANSSI enables new linkages between statistical and data sciences and research in academia, industry and government.

Strategic Priorities

1. Make Canada a global leader in multi-disciplinary research involving statistical science and data science
2. Advance Canada's research excellence in all fields by enabling collaborative research led by statistical and data scientists
3. Engage statistical and data scientists across the country in a vibrant community that promotes data-enabled learning in all spheres
4. Provide advanced training for multi-disciplinary collaborations involving statistical science and data science
5. Enable innovation and knowledge translation through collaborations with government and industry
6. Lead the development of a national data science strategy



Strategic Priorities, Actions, Target

Make Canada a global leader in multi-disciplinary research involving statistical science and data science

Canada has an excellent global reputation in statistical science: our researchers are internationally renowned, and Statistics Canada has long been considered one of the best government statistical agencies in the world. Five universities in Canada are in the top 50 in the QS rankings of statistics and operational research, and eight are in the top 100.

National efforts in multi-disciplinary research have been less prominent, at least in the close collaborations in science that we envision for CANSSI. Multi-disciplinary efforts in other countries are often seeded through special programs, for example at the National Science Foundation in the US, or through targeted initiatives, such as the Alan Turing Institute in the UK.

Where we are

Our Collaborative Research Teams (CRTs) all have international collaborators, and contribute to the global research enterprise. The members of our Scientific Advisory Committee, which reviews CRT proposals, are with two exceptions scientists from outside Canada. We have very strong links with the Statistical and Applied Mathematical Sciences Institute (SAMSI) and the National Institute for Statistical Sciences (NISS), both based in Research Triangle Park in North Carolina. We regularly sponsor invited paper sessions to the Joint Statistical Meetings. In 2017 the research of two CRTs was featured in an invited paper session at the annual meeting of the American Association for the Advancement of Science.

We are jointly sponsoring a SAMSI/CANSSI postdoctoral fellow in 2017 – 2019, and we have an agreement with SAMSI to sponsor undergraduate students in their undergraduate research workshops. The Distinguished Visitor Program brings internationally renowned researchers to Canada.

[Example sidebar: picture from Mills Flemming’s team and caption re the work with Anders Nielsen at DTU]

[Example sidebar: photo from a DVP lecture, or snap from Guttorp’s poster]

Going forward

We believe CANSSI has a model for enabling collaboration that is unique around the world, in statistical science. We must continue to raise international awareness of CANSSI and the CRT model, through reciprocal agreements and improved communications. We should leverage our individual researchers’ networks of contacts.



We should build our international profile through workshops and exchanges of postdoctoral fellows and graduate students.

In contrast, the world is well ahead of Canada in data science, with large initiatives having been established in many countries. CANSSI can play a pivotal role in developing large-scale data science efforts in Canada.

Action items

- Identify Canadian statistical scientists whose research interests overlap with upcoming thematic programs at SAMSI, the Newton Institute, and other international programs, and empower these to create activities in Canada associated with these programs
- Promote CANSSI through regular communications to international academic organizations, including societies such as the International Statistical Institute, the Royal Statistical Society, the Institute of Mathematical Statistics, and the American Statistical Association
- Use CANSSI's network to advertise and encourage applications to various visitor programs at the Alan Turing Institute
- Continue to encourage CRTs to apply for research workshops at the Banff International Research Station
- Explore opportunities to expand the number of BIRS-type workshops in statistical and data science
- Provide opportunities for trainees and early-career researchers to interact with global international experience in data science

Advance Canada's research excellence in all fields by enabling collaborative research with leadership from statistical and data scientists

Statistical science research is inherently multi-disciplinary. The foundations of statistical theory are continually assessed and revised in the light of experience and their relevance to the wide variety of contexts in which statistical considerations arise. Statistical researchers in academia are involved in collaborations with scientists, social scientists, humanists, government and industry, to a greater or lesser extent, throughout their careers. These interactions motivate new developments in statistical methods and theory, which then inform advances in application areas distinct from the ones in which they evolved. This ongoing interplay enriches the discipline and the profession.



Where we are

CANSSI was developed to both facilitate and elevate these collaborations, building on the earlier success of the National Program on Complex Data Structures --- the first national effort to have as its mandate fostering interdisciplinary research with statistical science leadership. Our flagship program of Collaborative Research Teams has built and strengthened research collaborations in fisheries stock assessment, wildfire prediction, financial modeling, environmental risk assessments, earth and atmospheric science, statistical genetics, neuro-imaging and public health.

The model we have developed, in which statistical leaders establish a research partnership with scientific leaders, is unique in Canada and possibly around the world. It emphasizes the co-creation of knowledge, the leveraging of Canada's statistical science expertise, and the multiplier effects achieved by embedding partnership into projects at their inception.

We are often asked by our colleagues in substantive fields to provide letters of support for grant applications, and partial support for workshops and conferences. We have to date responded positively to requests from colleagues in social science, digital humanities, medical physics, human genetics, applied mathematics and mathematical physics. Our view is that each of these represents a potential future collaboration.

[Example Sidebar: "I truly believe that it is only when statisticians can come to the table with funding, and keep HQP in their 'labs', that we can fully contribute".]

Going forward

We need to strengthen and expand the collaborative research model. CANSSI represents a substantial value added to a research program---partnerships across disciplines can be very difficult for an individual investigator to develop within an existing research grant. Our current funding requires that we emphasize fields in the mandate of the Natural Sciences and Engineering Research Council, but the model translates easily to the mandates of the Social Sciences and Humanities Research Council and the Canadian Institutes for Health Research.

Broadening our funding base will help to scale up this model well beyond our current reach. We aim to bring the power of our model and our expertise to the full domain of modern research in science, social science, humanities, and health and life science.

CANSSI will continue to have an essential role in the emerging field of data science. Statistical scientists have a unique perspective on questions around data quality and



provenance, on questions around inference from data, and on the reliability and precision of this inference. We will continue to work with the research communities in computer science and applied mathematics to develop new theoretical and practical approaches to the foundations of data science.

Action items

- Continue to dialogue with NSERC on the CTRMS program, which funds the network of institutes in mathematical and statistical sciences. Solicit support for this from partner organizations within CRTs and outside.
- Promote CANSSI with the President of CIHR and relevant Institute Directors, including the Institutes of Genetics, Health Services and Policy Research, and Population and Public Health to build on our investment in our network of Health Science Collaborating Centres
- Promote CANSSI with the President of SSHRC and cognate groups in social sciences and humanities, such as the Digital Humanities Summer Institute
- Build bridges to the Vector Institute, IVADO, and the Alberta Machine Intelligence Institute
- Participate actively with national groups working to develop the foundations of data science in Canada
- Review the CRT program to assess its strengths and weaknesses

[Example sidebar: list of supported CRTs to date]

[Engage statistical and data scientists across the country in a vibrant community that promotes data-enabled learning in all spheres](#)

Where we are

Academic statisticians work in all regions of the country, and many are somewhat isolated. By building national teams we help to provide critical mass and link isolated statisticians with established centres. It is a measure of our success at building community that we have 31 institutional members among departments of statistics, mathematics and statistics, or biostatistics, that are contributing financially to CANSSI.

Support for workshops and conferences helps to engage statistical scientists who are not directly involved in CRTs, and often serve as a platform for the development of new collaborations. The KickStart program was specifically designed to help isolated statisticians develop new collaborative directions.



In support of undergraduate training, we have supported two very successful workshops on statistical science education, and contributed funding to three data competitions. We have an agreement with SAMSI that reserves some places in their undergraduate research workshops for students from Canada supported by CANSSI.

[Example sidebar: quote from Redelmeier’s report on SAMSI workshop]

[Example sidebar: photo from the Sports Analytics meeting at SFU]

Going forward

As we build outwards towards data science, health science, and social science, we will need to find strategies to maintain this sense of community across a broader range of participants.

To date our efforts at outreach beyond academia, to, for example, secondary schools, and to professional statisticians working in government and industry, have been limited. The widespread interest in data-driven solutions should position CANSSI to be very influential in outreach. In 2017 we launched a week-long course on statistical methods for machine learning at the Digital Humanities Summer Institute; the course will be offered again at DHSI 2018.

Action items

- Partner with the Education Committee of the Statistical Society of Canada and the educational activities of the mathematical sciences institutes to provide resources to public schools on data science and statistical science
- Offer additional funding to CRTs to present their research to the general public and to local high schools
- Solicit and encourage CANSSI memberships from outside academic departments, including industry partners and relevant government agencies
- Build on outreach efforts of individual CRTs, some of which are quite substantial, to strengthen the CANSSI brand in government and industry
- Develop connections with CS-Can/Info-Can to build links between computer science and statistics

[Provide advanced training for multi-disciplinary collaborations involving statistical science and data science](#)



Learning to collaborate with scientists requires a considerable investment of time, a healthy dose of serendipity, and a knack for listening. By providing support for graduate student researchers and post-doctoral fellows that is targeted on collaborations, this learning can become an integral part of training the next generation of statistical and data scientists.

[Example sidebar: photo of students]

Where we are

CANSSI has several activities related to training the next generation. Collaborative Research Teams are required to use 60% of their funds for graduate students and postdoctoral fellows. Most teams have found that a postdoctoral fellow serves a key role in the team, often serving as the “glue” that keeps the group collaborating efficiently. Several teams have organized summer schools. Other teams have offered in term courses at two or more locations by webinar.

In response to requests from our community, we established a small national CANSSI postdoctoral fellowship program separate from the CRTs, which provides matching funding to researchers to enable them to fund a postdoctoral fellow for one year. In support of our interdisciplinary mission, we have responded where feasible to opportunities to partner with other organizations in support of postdoctoral fellowships. To date we have extended our postdoctoral program through partnership with the Statistics Laboratory at the Centre de recherches mathématiques, the Ontario Institute for Cancer Research, and the NSF-funded Statistical and Applied Mathematical Sciences Institute (SAMSI) in North Carolina, and the Postdoctoral Training Centre in Stochastics at PIMS.

In 2017 we launched a program of Health Science Collaborating Centres, designating seven groups across the country and providing modest seed funding to establish experiential training programs at the MSc and PhD level.

[Example sidebar: list of HSCCs]

We have three programs for undergraduate students: support to attend SAMSI undergraduate research workshops, support for students in statistics attending the MILA summer school on deep learning, and support for units leading a Datathon competition.

Going forward

We need to ensure that we can attract very strong postdoctoral researchers, both within CRTs and in our stand-alone and partnership programs. The pool of qualified



statistical scientists for this is not large, but as our reputation has grown we have seen a marked increase in quality of applicants. Postdoctoral fellowships are a more regular part of training in applied mathematics and in computer science, and strengthening our linkages there will expand our opportunities to attract world-leading candidates.

The strong response to our call for Health Science Collaborating Centres, with the emphasis on experiential learning, and the enthusiastic response from the Digital Humanities Summer Institute, suggests that there is considerable appetite for training programs of various lengths and at various levels. CANSSI is well-placed to offer an ongoing suite of training courses.

[Example Sidebar: “offering fellowships is good not only for the postdoc but also for faculty (especially younger ones) who would otherwise be unable to do this.”]

Action Items

- Review the capacity of the community to support an expanded postdoctoral fellowship program, including addressing the quality of the applicant pool
- Work with CRTs to coordinate summer training programs and explore opening these to graduate students outside of CRTs
- Work with the Health Science Collaborating Centres to build links with provincial and regional health authorities
- Develop a series of short training workshops on advanced topics in statistical and data science
- Work with our international partners to develop an exchange program for postdoctoral fellows

Enable innovation and knowledge translation through collaborations with government and industry

As in research, the demand for statistical expertise and collaboration in government and industry is always strong, both in research programs and in the ongoing demand for highly qualified statistical and data scientists.

Where we are

The ASA Datafest Competition and other similar datathons are excellent links to local companies, who sponsor the prizes, and send both mentors and judges to the competitions. Some individual CRTs have strong connections to relevant industries, and



several Mitacs internships: for example the CRT on dependence modeling via copulas has partnerships with the Banque de Montréal, the Banque Nationale du Canada, Desjardins Groupe d'assurances générales and Électricité de France. Our Industrial Innovation Committee has been working over the past two years jointly with the Industrial Innovation Platform of the mathematical sciences institutes on Industrial Problem Solving Workshops and associated training programs. We have excellent links with Statistics Canada, who used our network this year to attract several interns.

Going forward

Our efforts in innovation and knowledge translation should be enhanced by broadening our reach into data science, computer science, and applied mathematics. We can build on our experience with Statistics Canada to explore opportunities for collaboration with other government agencies, both federal and provincial. The experience of our CRTs can be leveraged for this: CRT 1 has an ongoing collaboration with Fisheries Canada, for example. We could have a program of industry-sponsored CRTs, with selected industry partners.

Action items

- Create affiliate memberships for industrial partners
- Track the placements of CANSSI-supported graduate students and PhDs in industry
- Explore the possibility of a formal relationship government agencies including Statistics Canada and Health Canada
- Identify CRT proposals or themes that could be suitable for industry sponsorship
- Build on the expertise of our Board members to contact key industry champions

Lead the development of a national data science strategy

Where we are

We have appointed a Data Sciences committee, to function similarly to the Health Sciences committee: assessing the opportunities for collaboration across Canada, promoting CANSSI as an essential partner in developing the field of data science, and highlighting opportunities for CANSSI to contribute to the research agenda in data science. We are preparing a position paper, jointly with the Statistical Society of Canada, clarifying the roles of statistical science in the rapidly changing field of data science. We are organizing an international workshop in fall 2018, in Toronto, on statistics and data science.



We are collaborating with the computer science community in a proposal for a Network of Centres of Excellence in Data Science.

Going forward

We should explore CRTs and other research activities joint with research groups in large data science companies---links with Canada's infrastructure in Artificial Intelligence should be useful for this. The community building that is enabled by the work on the NCE proposal in Data Science will lead to other opportunities for collaborations in data science.

Action items

- Request the Data Science committee to follow up on its 2017 report to the Board with recommendations for establishing a network of Data Science Collaborating Centres, modelled after the Health Science Collaborating Centres
- Partner with national data science programs to enhance specialized training through short courses
- Contribute leadership to the development of a national program and platform for data science research, recognizing that its considerations and concerns are now part of every field that uses data

[Example sidebar: snapshot of either CRM risk program or Big Data poster at Fields]