Working with IBM

When Premier Darrell Dexter announced last autumn that IBM would be taking over Nova Scotia’s SAP computer system, it represented a big opportunity for the Faculty of Computer Science. Mr. Dexter focused on the jobs to be created, as IBM takes over support for payroll, procurement, human resources and other information from government departments, district health authorities, school boards, housing authorities and some municipal bodies. The deal, however, also includes a forward-looking agreement between IBM and a consortium of universities in the province to develop training and research in the areas of analytics and optimization.

IBM is looking to the future, recognizing that their success depends upon finding people with the necessary skills in these rapidly growing fields and working with universities to meet this coming demand. The Dalhousie Faculty of Computer Science will be at the heart of this collaboration with our growing expertise in data analytics and in the relatively new field of “big data”. Big data refers to a data collection that is so large and so complex that it is difficult to process and digest using traditional database and data processing techniques. The world now produces huge amounts of data, and the rate of this is accelerating quickly. Ninety percent of the data that exists in the world today was produced only in the last two years.

One of the first Faculty initiatives will be to offer a new specialization in Data Science at the undergraduate level in September 2014, to go along with the three existing undergraduate specializations in Graphics, Gaming and Media, Artificial Intelligence and Intelligent Systems and Communication Technologies and Cyber Security. This may be a national first, as no other university in Canada currently offers such a specialization in data science. Graduates can expect to find their skills to be in hot demand, not only from companies in the IT sector, like IBM, but across a range of industries that contain large amounts of data.

IBM will help to build curriculum by sharing materials they have been developing, and by providing access to their personnel and case studies; and as part of their commitment to this program, they will help promote the Faculty’s courses worldwide.

Dalhousie’s Faculty of Computer Science is also growing in other areas to adapt to these changes and opportunities. Headed by our Tier 1 Canada
Working with IBM – big ideas and big data

Research Chair, Dr. Stan Matwin (see profile next), the Faculty has founded a new Institute for Big Data Analytics. This will become a national centre of excellence, supporting research, collaborating with industry and providing leadership in this developing field.

On the research level, IBM will provide in-kind support for designated projects and programs. New synergies between current computer science research issues and the big data research needs in other fields such as oceanography, health and business will hopefully be formed. IBM will provide software and equipment to support these courses and research. IBM in Nova Scotia is a big idea. It is more than a service contract, but rather a collaboration on a wide scale that will address the need for research and for highly skilled people as this field continues to grow. It’s a seed that will attract other companies to locate here, building a local ecology in expertise, research, industry and training that will have long term benefits to the province, its economy, the universities and the people of Nova Scotia.

The way that the term “Big Data” is thrown about these days might easily lead one to dismiss it as the latest buzzword, a transient obsession in a high-tech world which will quickly move on to the next thing. However, there is every reason to believe that “Big Data” will be with us for a while. It is a label for a problem that faces the widest range of human activities — the rapid accumulation of piles of data which we have yet to figure out how to fully deal with.

If you consider that there are 27 billion credit card transactions in the USA in one year, and then extrapolate that to the whole world, then add all the other types of financial transaction it quickly becomes apparent that the financial industries are generating a very big pile of data. A similar situation can be seen in every industry, government department or research program that generates and uses data.

Graduates will find themselves in a world that urgently needs their skills. The Globe and Mail tells us that the number of people employed in this field doubled in 2011 and will double again by 2015. The field currently has an unemployment rate of less than 1% and one study anticipates that, by 2018, there will be approximately 140,000 – 190,000 unfilled positions in data analytics in the USA. “Big Data”, it seems, is also about big opportunities.
What is your line of research?
I work in Machine Learning and Data Mining. Machine Learning is a research area in which a computer is given examples of something (e.g., what is and what isn’t an oil spill in a satellite image of the sea) and, from these examples, it learns how to classify or predict new examples of that “something” (e.g., to recognize oil spills in new, unseen images). This is an old idea, dating back to the 1950s, and it was part of the original Artificial Intelligence manifesto. Everybody agrees that learning is an inherent part of intelligence, but I like to see it more pragmatically. I am interested in the use of learning programs to learn practical things: to predict who in the emergency room will need hospitalization, to recognize oil spills, to categorize medical articles or to catch emerging trends in a political campaign or in public opinion.

Of particular interest for me is learning from text data: papers, blogs, tweets, notes, etc. I believe that such data calls for methods that take into account its linguistic character — we will have stronger methods if they understand the lexical, syntactic and semantic character of such data. This is the main topic of my Canada Research Chair here at Dal.

Data Mining, for me, is Machine Learning in the large. First, one is dealing with large data sets in millions of records and terabytes of volume. Second — in data mining — it is recognized that one spends most of their effort not in the “model building” phase, but instead in the data cleaning and data preparation phase (e.g., doing “attribute engineering”). In order to do this, the data miner must learn the basics of the domain from which the data is coming: they will have to create in their head fundamental “ontology” of that domain: what are the main entities and what are the relationships between those entities.

I am also interested in data privacy. I work on methods that make it hard, or practically impossible, to identify a given person in a dataset.

How did you get interested in that?
Well, it all started many years ago when I was involved in one of the early projects in Expert Systems (ES), a joint project with Cognos. At that time, we were trying to build an ES that would process (or assist in processing) government travel claims. I got to learn more than I ever wished about that “fascinating” topic! A question which arose was, how does one acquire rules which form the knowledge base of an expert system? Somebody suggested that I look at Machine Learning — indeed, one of its early goals was to replace the classical “Knowledge Acquisition” approach with learning the rules from examples. I went to spend a sabbatical with one of the
leading centres of Machine Learning at that time, George Mason University in Virginia, and I caught the bug. I liked the fact that Machine Learning was drawing on a variety of disciplines (AI, logic, databases and statistics) to build its tools. I also liked the fact that it was directly applicable almost everywhere. I am always interested in applications — they are an opportunity to learn about something completely new, from neuro-ophthalmology to forestry to electronic components (to name a few applications I was involved in). Applications also attract students and, last but not least, research funds. Done well, they often present a general research problem that can be shared with the community and initiate a new line of research. That has happened to our work on oil spill detection with R.C. Holte and M. Kubat that opened the active field of learning from imbalanced data.

My interest in data privacy is a little different. I am concerned about the fact that modern computers may become a tool that can be used to breach and violate people’s privacy easier and on a much larger scale than it was possible, say, 30 years ago. I believe that since the computer research community invented the tools that make it possible — databases, the internet, image and voice recognition, barcodes, etc., — it is then our moral obligation to at least think about tools that would make privacy easier and that would avoid many privacy-averse incidents.

What do you hope to achieve in the next five years?
I have several goals. First and foremost, I hope to create — together with colleagues from Dal — an active, dynamic centre of excellence in our joint field of research, which we call Big Data Analytics. We have recently created the Institute for Big Data Analytics to focus research on this area. The Institute will attract talent, ideas and applications, and will make Dalhousie a globally visible centre for this type of research. We’re getting a very powerful computer, IBM Netezza, a unique machine not only here but on campuses generally, which will provide an excellent infrastructure for Big Data applications.

I believe that since the computer research community invented the tools that make it possible — databases, the internet, image and voice recognition, barcodes, etc., — it is then our moral obligation to at least think about tools that would make privacy easier and that would avoid many privacy-averse incidents.

At the research level, I hope to make inroads into a linguistically informed but still scalable text model (“representation”). I want to complete several real-life, deployed applications of data and text mining techniques. I also want to continue with a start-up, Devera Logic, that I founded several years ago with colleagues in Ottawa in the area of computer security, and to bring it to a fruitful completion.

Who else is involved in this research?
Here at Dal there are several excellent researchers involved in this type of research. My closest collaborators in text analytics will be Dr. Vlado Keselj, Dr. Evangelos Milios and Dr. Mike Shep-

What attracts your interest outside your research area?
I am interested in current affairs and politics — I believe we have to be informed to influence decision makers on matters that concern us. I spend a lot of time reading (online) newspapers in at least three languages — English, French and Polish. I am also an avid reader of literature in these three languages. Classical music is my major hobby — I have a large CD collection, I go to concerts wherever I can, also during my frequent travel. I like hiking and swimming, but I do not do enough of that.
Art in motion: How “Patch” colourfully represents life in the Mona Campbell Building — a project by Dal CS student Stephen Kelly

Though the Mona Campbell Building is newer and greener than the rest of Dal’s buildings, in some ways it’s no different than the others: concrete and steel brought to life by the people who spend their days working, writing and thinking inside its walls.

Now, 18 glowing cube-shaped lanterns, hung over the stairwell just above the building’s third floor, are capturing that life and sharing it through colour and movement.

The art installation is called Patch, and it’s the work of Dal Computer Science PhD student Stephen Kelly. Its behaviours aren’t random: the cubes move and change colour based on data from the building’s environmental systems.

At the moment, their motion is based on the carbon dioxide levels in each of the building’s classrooms: the more students there are in the classrooms, the more the lanterns will rock back and forth. And as the building’s water usage spikes — pre- or post-class washroom breaks, perhaps — the cubes change from dark blue and begin to light up with yellows, purples and greens.

“I was looking for something that would feel alive,” explains Kelly, speaking to his inspiration for the design. “I wanted to represent the human activity in the building in a way that would bring to mind light.”

“The [building’s] green roof was also an inspiration — the way [the lanterns] move like blades of grass,” he adds, pointing at the plastic tubes that wave as the cubes start to sway.

Bridging talents

The project has had a long gestation: Kelly has been working on it in stages over three years, and the original idea goes back even further. It began during the planning stages for the building itself, when Peter Dykhuis, director of the Dalhousie Art Gallery, and Rochelle Owen, director of sustainability, identified an opportunity through the Canada Council for the Arts’ New Media Residency Grant to create a unique piece for what was to become the university’s greenest building.

Kelly was chosen for his unique combination of proven artistic talent — he was longlisted for the Sobey Art Award last year — and programming expertise. He worked with staff from the Office of Sustainability and Facilities Management to tap into the building’s data systems and make them talk with the installation.
Dykhuis, speaking at a reception on February 15th, said that Patch is an example of what’s possible, “when research interests in visual art and computer science combine to invent something new, something creative and artistic, that also has social relevance pertaining to visual data and a LEED-certified building.”

He also noted how appropriate it was to have a NSCAD University grad like Kelly bring an original, innovative art piece to the Mona Campbell Building. “This building is on the site of the campus of NSCAD from 1957 to 1974, where legendary artists created and exhibited groundbreaking new work,” said Dykhuis. “From a visual arts point of view in Halifax, this is sacred territory.”

A cross-discipline creation
Kelly, a full-time PhD student, expects to be on campus for at least a couple more years, during which time he will continue updating and building upon Patch. For example, he plans on incorporating other colours that will reflect changes in energy use or steam released by the building.

“I’m always interested in creating art that has a life of its own, that has an aesthetic or behaviour that reaches beyond my control,” he says.

Dykhuis and Kelly thanked the many individuals and groups who helped guide the project. They came from all corners of the university and beyond, demonstrating just how interdisciplinary the project was: Facilities Management, Dal Engineering, Dal Architecture, Dal Libraries, the College of Sustainability, the Office of Sustainability, Johnson Controls, the building’s architect (George Cotaras), NSCAD and many more.

“I just hope that it generates some stimulating conversations about the building,” says Kelly.
From academia to industry
How these partnerships can benefit FCS students

Recent Dalhousie PhD graduate, Marek Lipczak, has been sponsored as a Mitacs Elevate postdoctoral fellow at Halifax-based Oris4.

The Mitacs Elevate program provides matching funds to Canadian companies for hiring new PhD graduates. This way, small-to-medium enterprises (SME) can afford to hire highly skilled people who can significantly add to their research and development capacity, and ultimately contribute to their innovation-based success. Through Mitacs Elevate, new PhDs create and lead a major industrial research project while also establishing vital connections within the local business community.

Postdoctoral researchers will gain an appreciation of the economic impact potential of applied research. If they choose a career within industrial research, they will contribute to Canada’s competitiveness through the development of new products and services. If they choose an academic career within a Canadian university, their skills will be applied to strengthening the transfer of technology from academia to industry, and vice versa, an area which has been identified as weak in Canada compared to other industrialized countries.

A Mitacs Elevate fellowship is a partnership among the fellow, the industrial partner and an academic supervisor. The academic supervisor provides mentorship and research support to the fellow, while also providing research expertise to both the fellow and the industrial partner. The fellow’s time is shared between the organization’s site and the university, with a minimum of 50 per cent of the time spent within industry.

The ultimate goal of creating a conduit for knowledge transfer between university and industry is helped by engaging graduate students with research problems and helping to promote their research through publications at conferences and in academic and industry journals.

“I’m excited by the opportunity to do research in an industrial setting,” says Dr. Lipczak. “I would recommend both Mitacs programs—Elevate as well as Accelerate—to any graduate student interested in combining theory and practice in their thesis work.”

Dr. Lipczak’s next two years will be spent working with the Oris4 innovation team to enhance the intelligence in text mining of the company’s flagship product which is aimed at exploiting a gap in the Enterprise Content Management (ECM) market.

ECM is one of the fastest growing areas in enterprise software. It is a solution that keeps an organization’s documents and other unstructured content organized in a manner that people can easily find it. Oris4 is targeted at SMEs, research teams and individual knowledge workers. Dr. Lipczak’s thesis research in the field of tag recommendations in social bookmarking and tagging services was a natural fit for this team. “We recognized that Marek’s research, code quality and great attitude would be an excellent fit for where we are taking Oris4,” says Peter Hickey, President. “Oris4 is a technology that self-organizes information in an intelligent manner that allows users to find it quickly. Tagging is a task nobody likes and few do well,” he says.

“Marek’s work can significantly improve the process and experience for users.” His thesis work has potential application to the automatic tagging of text resources, leading to significant advances in searching small-scale document corpora, which will enhance the capabilities of Oris4 and contribute to its success in the marketplace.

“For me, it is especially interesting because thanks to the development of a Social Web providing large amounts of training data, we observe a great improvement in the applicability and accuracy of the machine learning algorithms that I am working on,” says Dr. Lipczak. “We have discovered a few areas where my experience from university can be applied to the development of Oris4, bringing it to a completely new level of intelligent interaction with the user. These improvements should lead to new, even more exciting research problems.”
From Brazil to Halifax
Dal welcomes students through Science Without Borders

When asked to talk about the differences between Canada and his home country of Brazil, Juliano Franz doesn’t go the obvious route. He doesn’t mention climate. He doesn’t identify cultural differences. He doesn’t compare Canada’s hockey obsession with Brazil’s soccer passion.

To Franz, it’s all about the chairs.

“Very comfortable chairs, especially in the Mona Campbell Building,” says Franz, a student at the Universidade Federal do Rio Grande do Sul (UFRGS) in Brazil who’s spending the next two semesters in Dal’s Computer Science program.

Franz came to Dal through the Science Without Borders program, a Brazilian government initiative that began in 2011 and aims to give out 100,000 scholarships over a four-year period to students wishing to study abroad.

Canadian universities will welcome about 12,000 of those students. This year, Dalhousie alone has 28 international students from Brazil.

Country comparisons
As one might expect, these visiting students have noticed things about their temporary home that locals take for granted. Franz’s comment about Dal’s classroom chairs, for example, aren’t a joke — or at least not entirely. In Brazil, he says, students don’t pay for post-secondary education, but that means some schools struggle with upkeep.

“You don’t have much money for infrastructure. Here, even the old buildings have been remodeled,” he says.

Franz’s countryman Hugo Yamamura came to Dal through CALDO, a consortium of four Canadian universities (University of Alberta, Universite Laval, Dalhousie and the University of Ottawa) that was one of the first groups to partner with Science Without Borders.

Yamamura, also a Computer Science student, is one of 24 Brazilian students who made their way to Dal via CALDO, which helps match international students with schools and programs appropriate to their interests and supports them through the application process.

Like Franz, Yamamura has picked up on some distinctly Canadian, or at least Haligonian, customs. For one thing, there’s the practice of cars stopping to allow people to cross in the middle of the street.

“You can go without any worries,” says Yamamura. “In Brazil, even at traffic lights you have to be careful.”

Embracing life at Dal
Yamamura and Franz have both embraced the full student experience at Dal, with Yamamura joining the chess club and Franz becoming a member of the squash club. They’ve found their fellow students and professors to be friendly and helpful, and they’ve adjusted to certain cultural nuances (“You have dinner too early here,” says Franz).

Yamamura says living in residence has helped them acclimatize to their new surroundings.

“Living on campus you have the opportunity to meet lots of people from all over Canada and even international students,” says Yamamura, who lives in Howe Hall.

Yamamura says Dal has also afforded him the opportunity to live near an ocean coastline. He’d never before visited a coast, even though Brazil has 7,367 kilometres of it.

“This was the first time I’d seen water in that proportion, not just a lake or rivers,” says Yamamura.

He and Franz have also enjoyed Haligonian sights such as Pier 21, the Seaport market and Point Pleasant Park.

Of course, Yamamura and Franz are here to learn, not just to explore a new country. Although both are in Computer Science, they’re taking different classes and have different goals for the future.

Yamamura is studying subjects like software development, interface design and microeconomics, while Franz is taking electronics, informatics and game design and development. Yamamura looks forward to a career in the private sector, while Franz sees himself as an academic.

A shared experience
Yamamura and Franz are different people, united by the shared experience of leaving Brazil for an adventure at Dal. They’ve made the adjustment to a new school and to life in Halifax, and they’re convinced they made the right decision.

That’s the message they have for all the future students who will make the trek from Brazil to Dal thanks to programs like CALDO and Science Without Borders.
This spring, the Faculty of Computer Science lost a long serving member when Dr. J Norman Scrimger retired.

Known as “Norm” (the J stands for James, in case you were wondering), he was a professor at the Technical University of Nova Scotia (TUNS) in 1997 when the Dalhousie Faculty of Computer Science was created through a merger of the relevant groups at TUNS and Dal. With computer science being a relatively new academic field, it is not unusual for professors of his generation to have started their academic careers in another area. Norm’s first passion was astronomy, gaining his PhD in 1979. He turned towards computer science after studying at TUNS in the mid 1980s.

Norm received the TUNS award for teaching excellence in 1992, the Srini Award for Teaching Excellence in 2003 and the Dalhousie Outstanding Academic Advisor Award in 2004. Norm has continuously been recognized for his unique ability to connect with students, to engage with his subject and to make sometimes difficult materials comprehensible. Norm also served as Dean from 2006 to 2007 and was the Director of Informatics until his retirement.

Over coffee in his office, Norm was prompted to look back and reflect on how things have changed since the Faculty of Computer Science began. Pointing at some out-of-date books on his bookshelf on Pascal and COBOL, he made the point that the fundamental principles of computer science really do not actually change so fast, and are still taught today pretty much as they were 16 years ago.

Of course, new fields — like network security — arise as computer scientists address the changing needs of industry and society. The dramatic increases in computer memory and processor power means that many more people can do their research on desktop computers, rather than using a central server, or “mainframe” as it used to be called. The rise of Internet has also brought its own changes to the way we communicate, teach, study and do research. Norm recalled sending an email in the early days to a collaborator in Dartmouth and noted with amusement how it had to be routed through the main NORAD server in Boulder, Colorado in order to get there. So although some teaching principles have remained the same, every year we have seen advancements that have altered the way we teach these principles.

Our facilities have also grown over the years. In the beginning, the new Faculty was housed in the Maritime Centre, at the bottom of Spring Garden Road, and only moved into its own building on University Avenue in 1999. Two years later, the fourth floor of the building was finally finished and, today, that floor is full of graduate student research groups, faculty and post-doc offices, a conference seminar space and a fifth teaching lab. Only two years ago, we expanded — yet again — to occupy the fourth floor of the new Mona Campbell Building, a space dedicated entirely to computer science research.

Norm pointed out that the makeup of the student body has also been slowly evolving. Today, many more students see the relevance of computer science to a variety of different career paths and come to us with a range of backgrounds and goals. In earlier days, the student body was made up more of committed computer enthusiasts and “hackers”.

We, now, also have a larger component of international students, creating a growing cultural and linguistic diversity in the Faculty.

Norm has always emphasized the human connection in computer science; his fondest memories are of interacting with students in the classroom, of convocations, celebrating student achievements and meeting families. His belief in the importance of the human connection was further confirmed when, in his role as Dean, he attended an alumni meeting in Ottawa. Many people there came to the event entirely for the personal contact — to connect with a face from their past, someone who helped them through the challenges of their student days.

Although Norm has handed in the key to his corner office he will still come back to Dalhousie. He plans to use his time to take courses in his other areas of interest; music and languages (human languages, that is, not computer languages!). He also plans to continue with his choral singing activities and to revisit his interests in photography and astro-photography, going back to his early interest in astronomy — where it all started for Norm.
An app-etizing contender

Food software wins Dal mobile app competition

It’s an app that works for dessert too.

Of course, in this case, “app” is short for application, but “DalEats,” which won Dal’s second-annual mobile app competition, certainly appealed to the foodies who dropped by the Goldberg Computer Science Building on March 27th.

The idea of the competition, hosted by the Faculty of Computer Science, is to challenge some of Dal’s best student programmers and developers by having them develop a functioning mobile app in less than a month. Specifically, the students were asked to design an app to benefit some portion of the Dal community, be it current or prospective students, faculty and staff, alumni or the surrounding community.

DalEats, developed by CS student Casey Yu, is a food app that allows students to track their dining on campus.

“It would include all the dining hall menus, as well as all the health information about the food,” she explained. “You can even subscribe to your favourite foods and get alerts when they’re going to be served.”

The app was one of many that impressed the judges, whose ranks included Vice-President Academic and Provost Carolyn Watters, Vice-President Finance and Administration Ken Burt and Assistant Vice-President Student Academic Success Services Meri Kim Oliver.

“You really blew everyone away,” said Dr. Watters. “Just a terrific set of apps made in a very short period of time.”

Making Yu’s win all the more impressive is that while many of the entries were by a team of students, hers was a solo effort. For her work, she received the competition’s top prize of $750.

Finance and Administration Ken Burt and Assistant Vice-President Student Academic Success Services Meri Kim Oliver.

Above: some screenshots of Casey’s winning “DalEats” food app
Class Notes
Please send in any news, announcements, events or ‘things you’d like to see’ to Allison.kincade@dal.ca. Stay connected with us online via Twitter.com/dalfcs and Facebook.com/dalfcs. We’d love to hear from you!

The Faculty of Computer Science has a new look online! Check us out at: cs.dal.ca

Contributors:
Allison Kincade, David Langstroth, Ryan McNutt, Stan Matwin, Michael Shepherd, Matt Semansky, Nick Pearce, Danny Abriel, Dave MacDonald, Blaine Wood