Data Day 2018: Giving voice to your research

May 2, 2018, 1966 Reading Room, Douglas Library

8:30 – 9:00  Registration – Coffee and Tea
9:00 – 9:15  Welcome and Opening Remarks
Martha Whitehead, Vice-Provost (Digital Planning) and University Librarian
9:15 - 9:25  Poll Everywhere Introduction
Shaun Cahill, Information Technology Services
9:25 – 10:40 Services and Infrastructure Panel
Panel Moderator – Karina McInnis, Executive Director, University Research Services

• Data Day 2018
  Jennifer Doyle, Chief Information Officer & Associate Vice Principal of Information
  Jennifer will describe how she understands her role in supporting research and researchers at Queen’s, discuss services available for researchers through ITS, and identify initiatives that are underway to address service gaps and opportunities.

• The Centre for Advanced Computing (CAC) Overview
  Dr. Ken Edgecombe, Director, Analytics Development Hub, Centre for Advanced Computing (CAC)
  The CAC serves Queen’s by providing advanced research computing, including hosting and storage solutions in a high-availability, high-security (PHIPA compliant) data centre. The CAC also offers analytics development and large data management services.

• How and Why the Queen’s University Research Ethics Boards are Keeping Research Data Safe
  Jennifer Couture, Ethics Compliance Advisor, University Research Services
  The Queen’s University General Research Ethics Board (GREB) and the Queen’s University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board (HSREB) abide by the principles outlined in the latest edition of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2 2014). The REBs are responsible for reviewing research proposals to ensure compliance with applicable regulations, and to ensure that the following principles are respected for research participants: Respect for Persons, Concern for Welfare, and Justice. This session will cover strategies for promoting the protection of data, and it will summarize why we need REBs to ensure that the rights and wellbeing of research participants are respected.

• Open Scholarship Services at the Queen’s University Library
  Courtney Matthews, Head, Open Scholarship Services
  Open Scholarship Services is a division of the Queen’s University Library that provides a suite of integrated services supporting the open sharing of digital research and scholarship including research data. This presentation will provide an overview of the local, regional, and national resources that underlie our integrated services and briefly discuss a training program that will guide researchers to manage their outputs openly and throughout the research lifecycle.
• **Data at Scholars Portal: Dataverse and other shared platforms for discovery, access, deposit, and re-use**

Meghan Goodchild, Research Data Management Systems Librarian, QUL and Scholars Portal

Scholars Portal, a service of the Ontario Council of University Libraries (OCUL), provides shared technological infrastructure and collections for university libraries in the province. This presentation will provide an introduction to platforms that allow users to explore and download geospatial datasets (Scholars GeoPortal) and social science/polling data (ODESI). Additionally, an overview of the openly accessible repository platform Dataverse will illustrate features related to depositing, sharing, citing, and exploring research data.

10:40-10:55  Coffee and Conversation
10:55-12:00  Research Experience Panel

Panel Moderator – Sarah Bartlett, Library Intern, Open Scholarship Services

• **Advanced Analytics at SCENE**

Dr. Stephen Thomas, Smith School of Business

SCENE is a loyalty program for movie/entertainment lovers. During the course of its operations, SCENE collects a massive volume of member data. Two particular data sets of interest include member demographics and member transactions. In this research, we aim to increase member engagement and recruit new members by using supervised and unsupervised machine learning techniques to uncover patterns of behaviour in the data, and make predictions about future member behaviour.

• **The Perils and Promise of Global Data Collection**

Andrew Coombs, Doctoral Student, Assessment & Evaluation Group, Faculty of Education

Everyday, classroom teachers use data to make decisions about what to teach and next steps for their students’ learning. Classroom level data is generated from various assessment tasks including traditional forms of assessment such as test, quizzes, essays, and reports, as well as more formative assessments such as conversations, student questions, and exit slips. There is no shortage of evidence to suggest that teachers’ classroom assessment practices have a profound effect on student learning and their achievement. However, due to teachers’ different approaches to classroom assessment practices, students experience learning cultures that value and use assessments in significantly different ways. My doctoral research examines how and why teachers’ approaches to assessment vary, primarily through the use of the Approaches to Classroom Assessment Inventory (ACAI).

Designed by Dr. Christopher DeLuca (Faculty of Education, Queen’s University) and his research team, the ACAI presents teachers with five classroom assessment scenarios (e.g., You discover that one of your students has plagiarized some of his/her/their assignment). Participants then decide how likely they are to complete 12 actions (e.g., I would grade aspects of student work that are original and deduct points for the plagiarized sections) on a six-point scale. At the end of the ACAI, teachers receive a personalized assessment profile that explains his/her/their approach to assessment across four dimensions (i.e., purpose, process, fairness, and theory). This profile can be used to guide further professional development.
In 2016, we administered the ACAI to a relatively small sample of teacher candidates and classroom teachers (n=404). Since then, we have conducted a series of studies with ever increasing sample sizes, diversified our studied populations, and integrated a variety of other instruments (e.g., Dweck Mindset Scale, Big Five Personality Traits, Self-Efficacy Scale). Furthermore, we have established a network of international partnerships in Australia, New Zealand, the United States, Great Britain, Germany and Canada with the aim of improving the development and support for teachers’ classroom assessment practices.

As the scope of the ACAI rapidly expands, the research team has faced several data management challenges. The issue of most immediate concern is the management of multiple streams of data across an international dataset. Since the ACAI is almost exclusively accessed through an online platform, this requires a great deal of planning and coordination between researchers to ensure their data can be easily identified, collected, and examined. Additionally, the centralization of raw data on a single platform has proven to have both benefits and limitations for such a diverse collection of global researchers. Furthermore, with the shift towards more internationally focused studies, the items within the ACAI have required periodic reconstruction including item type (e.g., 6-point scale, ranking) and word choice/phrasing to ensure their cultural relevance. While this has proven to be a resource-intensive effort, it has allowed the research team to address ever more complex research questions.

With a growing collection of researchers interested in using the ACAI to examine teachers’ approaches to assessment, new data management issues will continue to arise. However, in gathering more complex data from a diversity of teachers, we are better able explain how and why teachers’ approaches to assessment vary to better support student learning and achievement.

- **A closer look at the scalability of spatiotemporal data in geological sciences and engineering**

  Danielle Beaulne, Fotopolous Lab, Department of Geological Sciences and Geological Engineering

  Big data has fundamentally transformed the spatiotemporal scales at which we study our world. Not only are we looking broader across cities, countries and continents, but we are also looking deeper by extracting higher resolution data and with greater frequency. In the geodesy and geophysics lab at Queen’s University, research is focused on collecting data from across heterogeneous spatial and temporal scales in order to understand the level of scalability between them. Fundamental questions include - do patterns observed at the micro- or meso-scale amplify to produce the patterns we observe at macroscales? Do processes that we monitor over short time periods extend to generate the processes that we observe over longer temporal scales? Key components of this research include promoting the accessibility of geoscientific data and advancing the integration of big data into the geological sciences. To investigate these relationships, we employ an extensive array of information ranging from hand samples, gravity, magnetometry, seismic, SAR, optical, LiDAR, laser imaging and genetic data collected from the lab, field surveys, UAV’s,
airplanes and satellites. Analyzing these data require an equivalently diverse set of software and services available at Queen's University including the geophysics high performance computing lab and the cluster hosted by Queen’s Center for Advances Computing.

- **Using text corpora to study language**
  Anastasia Riehl, Director, Strathy Language Unit
  Large digital corpora of written texts or transcribed speech are becoming increasingly common tools in the study of language. In this presentation, I introduce two different types of corpus projects based at the Strathy Language Unit: the Strathy Corpus of Canadian English and the Wolfe Island English Corpus

12:00 – 12:20 **Research Data Management (RDM) Survey Summary of Findings: Navigating the RDM Landscape at Queen’s University and beyond**
Francine Berish, Geospatial Data Librarian

12:20 – 12:50 **Research Data Management in Canada – Portage Update**
Jeff Moon, Director, Portage Network
Portage is a national, library-based research data management network that coalesces initiatives in research data management to build capacity and to coordinate activities better.

12:50 – 1:00 **Networking Lunch Introduction**
Heather McMullen, Associate University Librarian

1:00 **Networking Lunch**