THE UNIVERSITY OF BRITISH COLUMBIA

Office of the Vice-President, Research & Innovation

UBC Research Data Management (RDM) Fall Series 2020

Post-Event Summary Report

Submitted by Nick Rochlin on behalf of RDM Fall Series Team

Event Summary

From September 28 - October 16, 2020, the *UBC RDM 2020 Fall Series* (hereinafter "Fall Series") was hosted collaboratively by the UBC Library, UBC Advanced Research Computing (ARC), the UBC Centre for Scholarly Communications (CSC), and the UBC Okanagan Library. The virtual Fall Series offered practical skills and guidance regarding research data management, and aimed to provide an overview of the entire research data lifecycle.

Courses in the series were taught by UBC Librarians from both Vancouver and Okanagan campuses, as well as members of UBC ARC. Most courses had a corresponding 2-hour Office Hour Session on the Friday morning of their respective weeks, encouraging attendees to bring questions specific to their research that would be too detailed to ask during the workshop sessions. The final session was a panel discussion with 2 UBC researchers and 1 staff member from the Centre for Brain Health, discussing the importance of data management from personal experiences. The courses covered the following:

- Introduction to Research Data Management (+ Office Hours)
- First Steps in Data Management Planning (+ Office Hours)
- Introduction to the DMP Assistant (+ Office Hours)
- Re-Introduction to the Data Life Cycle (+ Office Hours)
- Introduction to the Open Science Framework
- Introduction to Data Repositories (+ Office Hours)
- Introduction to Dataverse (+ Office Hours)
- Data Discovery Using FRDR and Globus (+ Office Hours)
- Introduction to Sensitive Data and De-Identification (+ Office Hours)
- Introduction to REDCap
- Data Horror Stories & Open Panel with UBC Researchers

Highlights

- 365 unique registrants, including:
    - 223 from UBC Vancouver
    - 89 from UBC Okanagan
- 684 total attendees
- Over 13 hours of sessions taught
- 16 hours of support via office hours
- Faculty of Medicine as top registered faculty
- Based on the goals and objectives, as well as large attendance numbers in an uncertain virtual environment, this event was considered a great success

Table of Contents

Introduction

This report provides a post-session summary of the three-week long *UBC RDM 2020 Fall Series* (hereinafter "Fall Series"), hosted by the UBC Library, UBC Advanced Research Computing (ARC), the Centre for Scholarly Communications (CSC), and the UBC Okanagan Library. The event was attended by UBC students, staff, faculty, and non-UBC individuals from across Canada and internationally. With 11 workshop sessions and 8 office hour sessions, there were nearly 13 hours of instruction and 16 hours of support, providing attendees with fundamental knowledge of research data management skills, tools, and practices.

Background and Rationale

For the past several years, Research Data Management (RDM) has been a growing focus for researchers and research funders, reflecting a cultural shift in research that prioritizes reproducibility and replicability, collaboration, and openness. While COVID-19 has postponed the launch of the Tri-Agency Research Data Management Policy (see Draft Policy for reference), it is vital that the education and training of RDM principles continues and evolves to drive this cultural shift, and to ensure that this instruction is not only based around compliance with upcoming requirements, but also build leaders and innovators in best RDM practices.

With the shift to a virtual work environment, an event that otherwise would have been isolated to a single physical location was opened up to a broader UBC and Canadian audience. This also provided an opportunity for collaboration across four distinct portfolios and both UBC campuses, which was a great first step in coordinating services and maximizing subject matter expertise on an institutional level.

Goals and Objectives

The primary purpose of this series was to provide training on RDM best practices beyond a general overview, and to build knowledge of tools and skills needed throughout the entire research data lifecycle, focusing on graduate students and early career researchers as our audience. As such, a mock research project was created for the series, containing a project description, a data management plan (DMP), datasets, and readme files. The mock project provided a practical grounding of the concepts covered in each session, and helped to tie the content from all sessions together.

Another objective of the series was to construct a baseline program and materials which could be further expanded and adapted. When the RDM Series team first started planning this event, the intention was to create a program that could run annually and continuously be tightened (please see the "Lessons Learned & Future Directions" section for plans of future RDM Series). The materials created for the series have also been placed in an Open Science Framework (OSF) wiki that is open for public use and adaptation, to help further the state of RDM training across Canada and beyond (OSF RDM Series).

Courses/Program

Among all courses offered throughout the Fall Series, the most popular was the “Introduction to RDM”, followed by “First Steps in Data Management Planning”. It is difficult to determine whether there was a greater interest in these introductory concepts, or if scheduling conflicts and other factors played into the popularity of the first two sessions. Regardless, while attendance numbers varied across sessions, the attendance rate was quite strong throughout the entire series with a mean value of 61%.

Each week of the RDM Series had its own theme, and the ordering of sessions and content began with broad/introductory topics and progressed to more specific topics, and also followed the temporal order of the steps in the research data lifecycle. The mock research project was utilized across all three weeks of the series to create cohesion and connectedness between the sessions and the weeks. Below is a summary of the weeks and their sessions

Week 1 - Introduction to RDM

- Introduction to RDM
- First Steps in Data Management Planning
- Introduction to the DMP Assistant
- Re-Introduction to the Data Life Cycle

Week 2 - Data Repositories

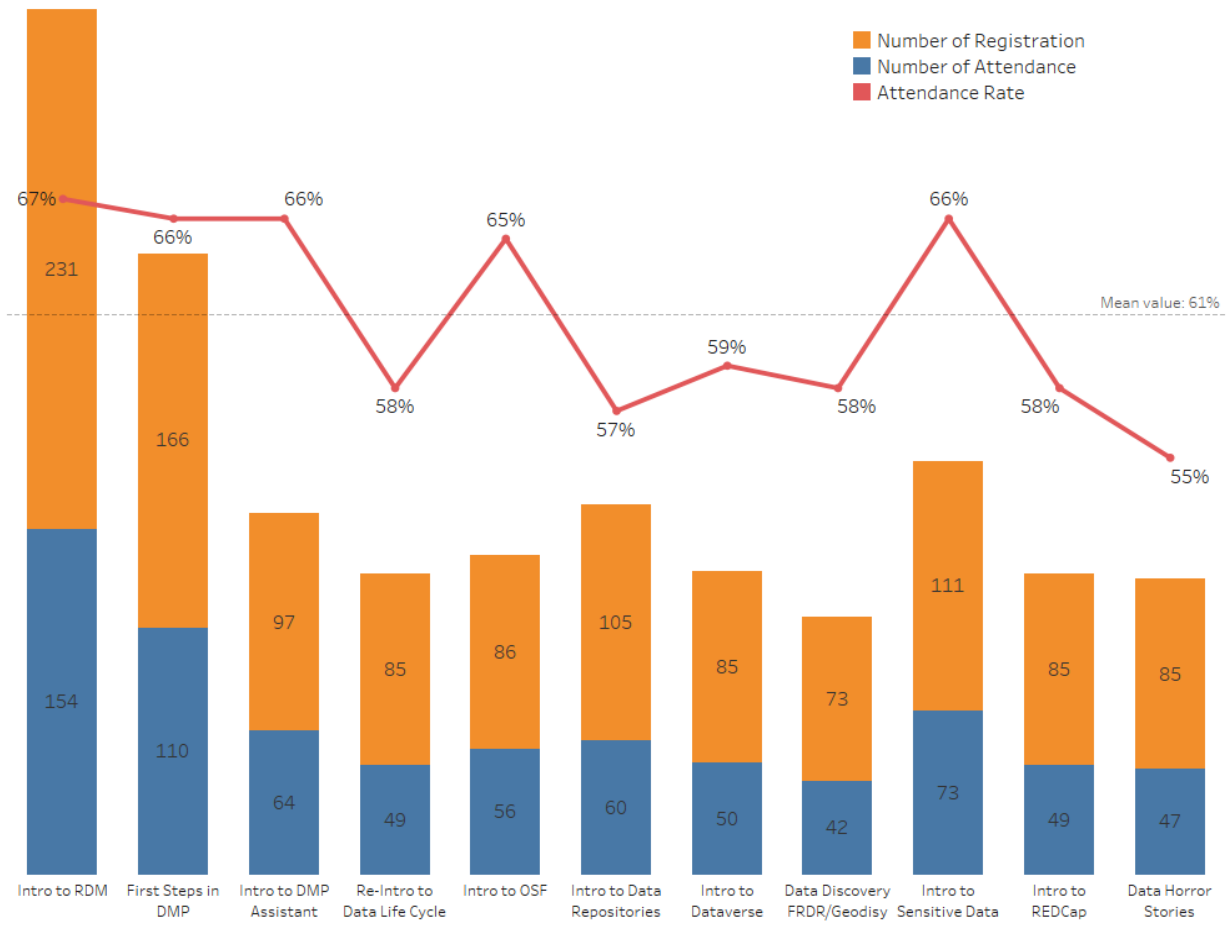
- Introduction to the Open Science Framework
- Introduction to Data Repositories
- Introduction to Dataverse
- Data Discovery and Deposit Using FRDR, Geodisy, and Globus

Week 3 - Sensitive Data

- Introduction to Sensitive Data and De-Identification
- Introduction to REDCap
- Data Horror Stories and Open Panel with UBC Researchers

Although it was not the session with the highest attendance, the “Data Horror Stories & Open Panel with UBC Researchers” session provided an opportunity for attendees to hear firsthand experiences, recommendations, and best practices of RDM from Dr. Jason Pither, Dr. Fiona McDonald, and Jeffrey LeDue. This session gave attendees an opportunity to ask questions to the panel which resulted in a number of interesting and thought provoking discussions.

Table 1: Attendance Rate



Attendance

As the Fall Series was entirely virtual, the decision was made to accommodate as many attendees as were interested. Promotion for this event was done through the Research Commons' website and calendar, the ARC website, all four partners' communication channels, as well as direct outreach to faculties, departments, and points-of-contact. Very close to the beginning of the Fall Series it was decided that the sessions should be open to a broader Canadian audience, and communications were sent out to a national listserv.

When looking at the attendance numbers by faculty (see Table 2), "Medicine" represents a strong majority (includes numbers from the Faculty of Medicine and Southern Medical Program), which may indicate a desire for Medicine-specific RDM workshops in the future.

Table 2: Unique Registration Breakdown by Faculty Affiliation (Total 365)

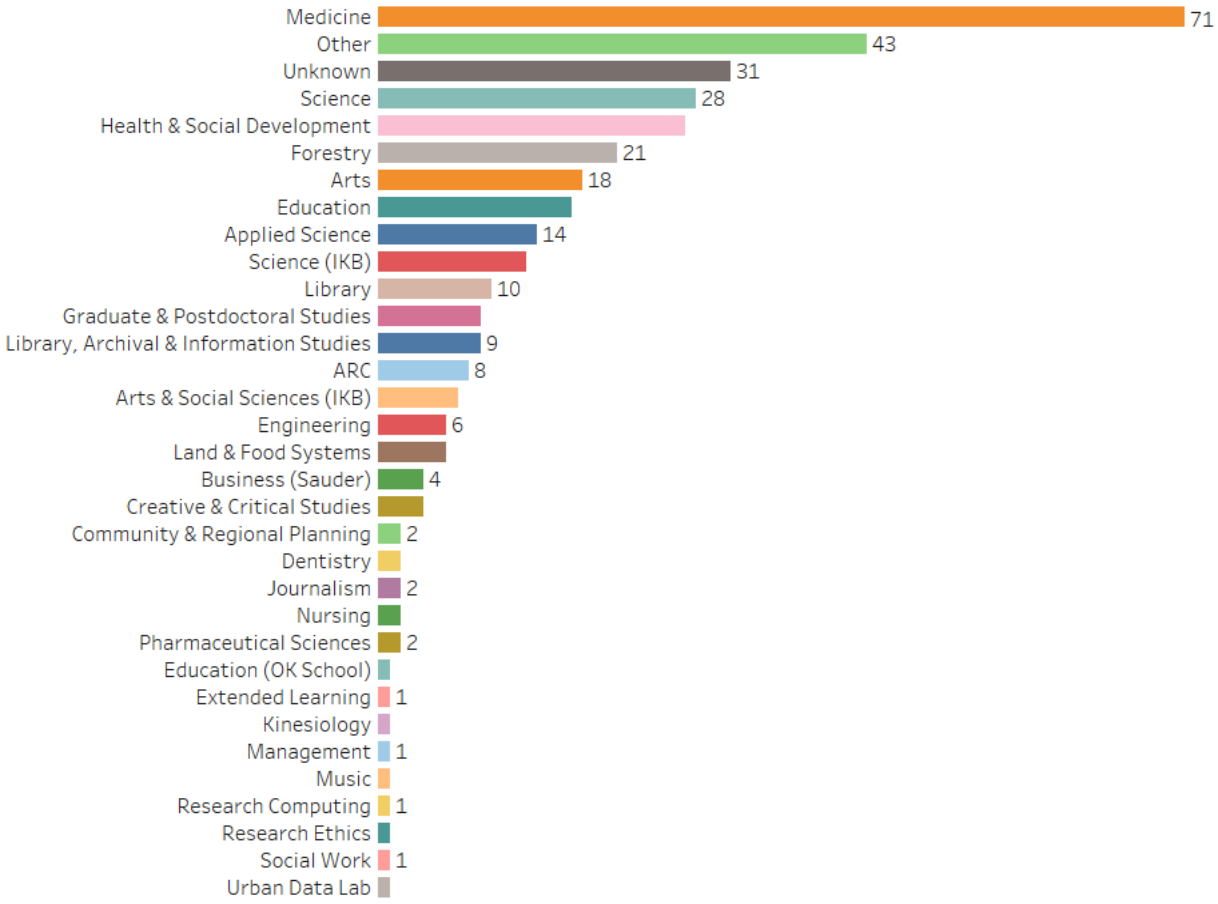
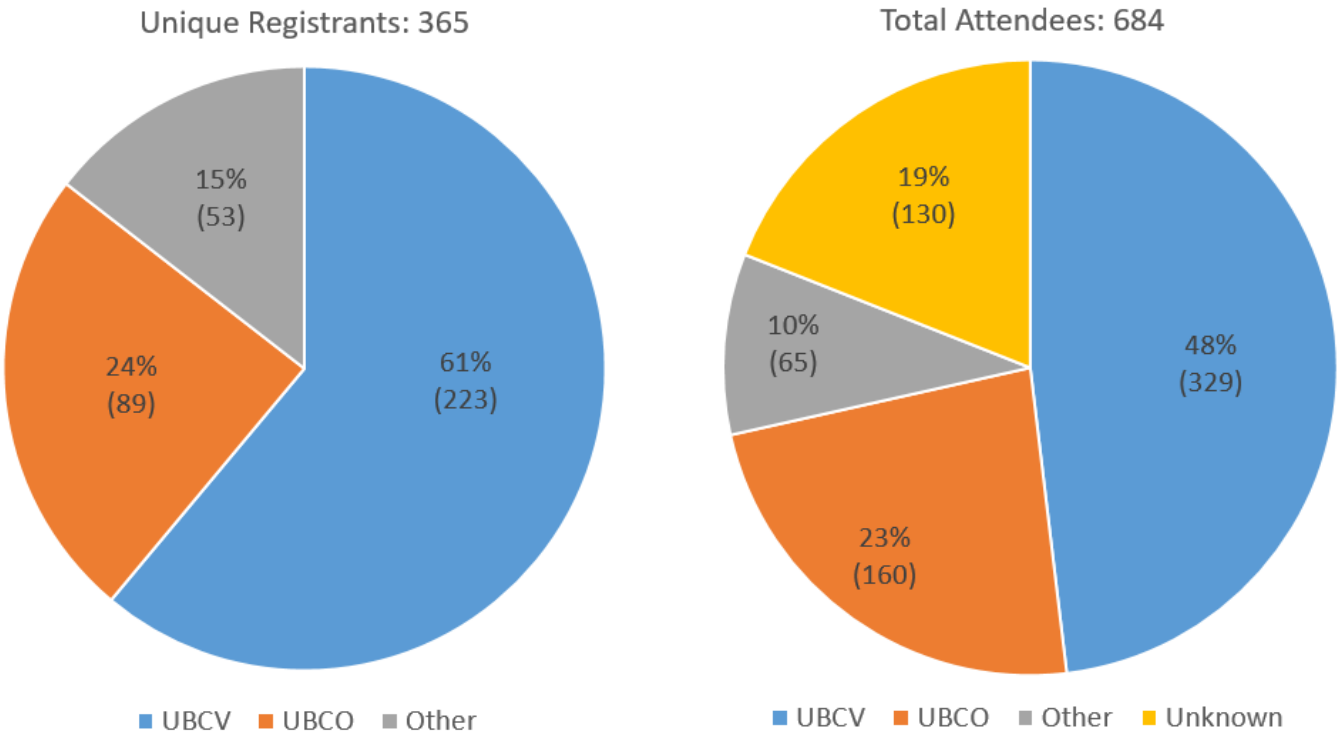


Table 3: Word cloud of non-UBC attendees (53 total)

Given the virtual environment, we had the opportunity to launch the Fall Series as a truly institutional event, and both campuses showed keen registration and attendance numbers (see Table 4).There is an “Other” category that represents those with affiliations outside of UBC, which can be seen in detail in Tables 3 and 4. There is also an “Unknown” category, which represents a null value in the affiliation field.

Table 4. Unique Registrants and Total Attendees by Campus Affiliation

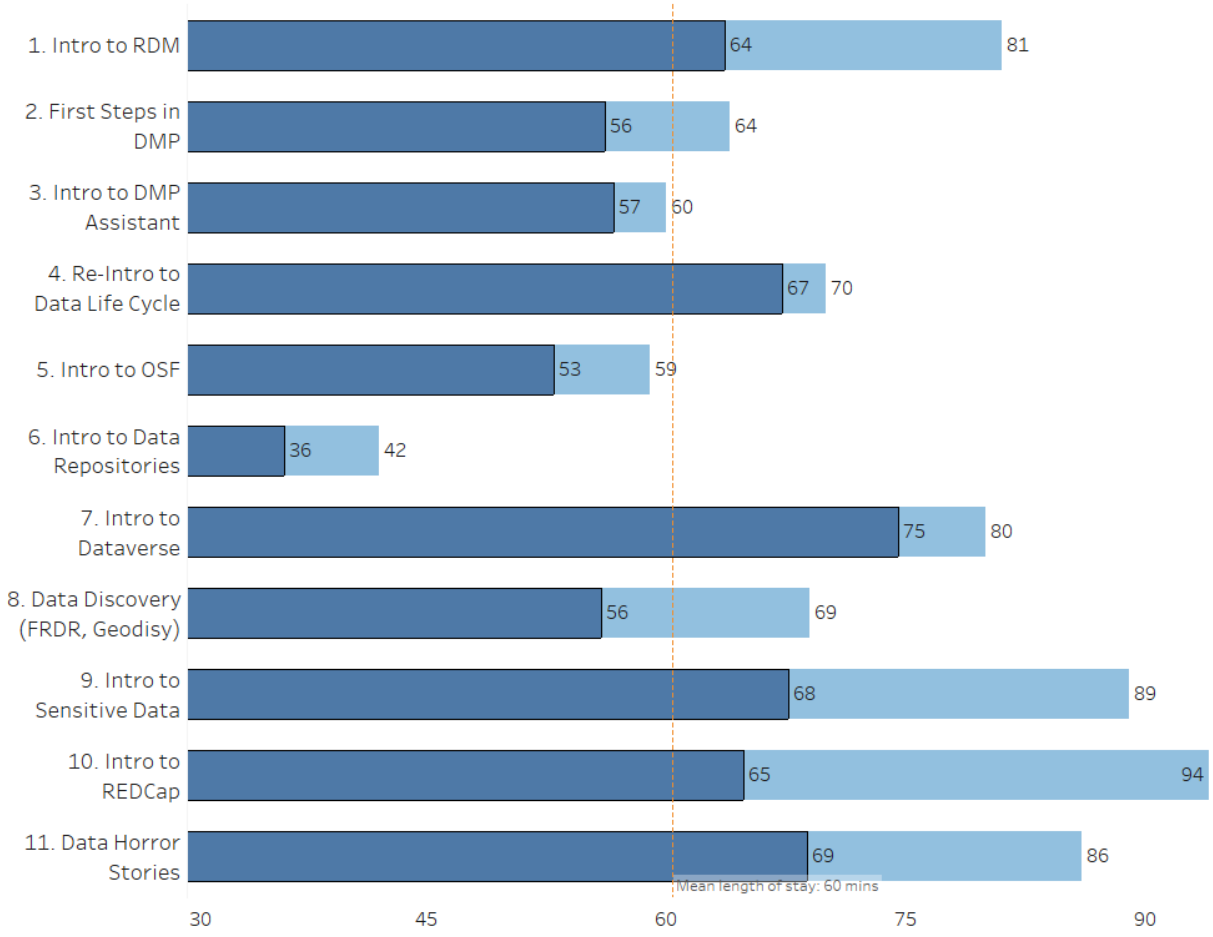


Okanagan Context

The RDM Series was a true collaboration across UBC institutionally. In this collaboration, it is worth noting the prominent role that the Okanagan campus played. There were significant contributions from ARC (Okanagan campus), the Okanagan Library, as well as the Centre for Scholarly Communications in the organization and execution of the event. Below are some Okanagan-specific highlights:

- 5 of 11 sessions featured a presenter from the Okanagan campus
- 2 of 3 researchers from the “Data Horror Stories” session were from the Okanagan
- Relative to its size (1/10 of the grad students and faculty of the Vancouver campus, source), Okanagan registration and attendance numbers were very strong (see Table 4)

Table 5: Average Length of Stay (minutes)



Effort to Create Series

Planning

In order to prepare and execute this event, the Fall Series team began meeting bi-weekly in mid-June to begin planning. Many of the early conversations were geared towards developing an innovative and practical program, and several draft programs were created before arriving at a finalized product. As the goal was to develop a series of sessions that were to flow from one to the next, and represent the unique stages of the research data life cycle, it was decided that creating a mock research project that would be used throughout the Series, would be an effective way to create a cohesive program. In wanting to create sessions that were relevant to as broad of an audience as possible, we chose the topic of Okanagan apple production, and created a project description, production dataset (raw and clean), worker satisfaction survey data (raw and clean), a data management plan (DMP) using the DMP Assistant, metadata for deposit, as well as readme files for the datasets (all can be found is the OSF wiki).

While one of the main goals of creating this content was to develop a high-quality event, another primary objective was to develop training materials that are openly shared for reuse and adaptation. As RDM is becoming an increasingly important issue in the Canadian research landscape, it's important that we contribute to this training ecosystem on a national level. This will not only help in positioning UBC as a national leader in RDM training, but will also allow for other institutions to adapt and expand on our materials, and drive the innovation and growth of RDM training.

Logistics

Significant time was dedicated to arranging the logistical aspects of the sessions, including such tasks as:

- Working with UBC IT to utilize Zoom webinar function
- Scheduling helpers for each session
- Arranging the distribution of session links
- Conducting dry-runs with all presenters to ensure technical aspects of the presentations worked
- Editing the recordings of all but 1 sessions (research panel was not recorded), and ensuring presenters' consent to post
- Ensuring OSF pages for each session contained:
    - Slides
    - Supplemental materials
    - Instructions and session information for helpers
    - Recording (posted within 2 days of session)

Promotion

Event was promoted on the Research Commons and ARC website, as well as through an extensive amount of newsletters and emails. For a full list of communications channels, as well as the blurb and image used for promotions, see Appendix A.

Feedback

After each session, registrants were sent a follow-up survey asking for written feedback using the Research Commons' template survey. Of 684 total attendees, there were only 34 submitted surveys. While these numbers are quite low, there were a few concepts to note.

"One thing you loved"

Common themes in for this question included:

- The interactive segments of the sessions
- The presenters and session delivery
- Active and thorough engagement with Q&A

"One thing you would change"

Of the 34 responses, 14 wrote a variation of "nothing" or left that field blank.
While there weren't any common themes for this question, some noteworthy feedback included:

- The 1.5 hour time slot is difficult for scheduling and attention span, 1 hour is easier
- Desire for domain-specific segments, as not everything presented was relevant to all research
- More explanation of purpose behind RDM best practices rather than presenting it as compliance oriented

Lessons Learned and Future Directions

Overall the planning and delivery of the series went very well, and the feedback we received was largely positive. One piece of feedback that was received verbally at the end of a session, and that was further elaborated in a follow-up meeting, was that it was difficult to see the connections between the sessions and there was a sense of information overload. While it will take some discussion and creative thinking to develop a more seamless model for the series, there are a few considerations that may help guide this process.

Condensed Delivery

One consideration is to hold future series in the summer, and to have content hosted in half or full-day formats to allow for more seamless connections between sessions, and for concepts to remain fresh in people's minds.

More than Just the "Book-Ends"

Most of the content covered in the Fall Series was situated around the planning and deposition of data, and did not delve into the collection, cleaning, and analysis of data, which are key components of data management. It may be worth including these elements in upcoming series as a way to illustrate a more complete image of the data lifecycle and to create more cohesion around the planning and deposition processes. As ARC hosts an annual Research Computing Summer School, and the Research Commons has regular sessions on research computing topics, bringing these more closely together may prove fruitful in delivering more comprehensive and research-focused training.

Discipline / Concept-Specific Sessions

One of the objectives when developing the series was to make the content as discipline agnostic as possible, with the goal of making the sessions broadly accessible and valuable. However, in doing this we unavoidably ran into the issue of not being specific enough or missing key concepts or practices of certain research types. A possible avenue for future series would be to develop sessions, or forked series, tailored for specific research communities or types of research.

Office Hours

While the office hours weren't as well attended as we had hoped, the idea of hosting virtual drop-in sessions is still one that we believe is valuable and worth pursuing. With that said, hosting separate office hours for each session was perhaps not the best mode of delivery. There were several occurrences of people showing up to the sessions with no questions of their own, and wanting to listen to the questions and discussions of others. Because the sessions were happening in separate rooms, and each session had to be registered for in advance, this made it challenging to host a discussion for open listening. In the future, hosting a single drop-in session will be better suited for our audience. In addition, using the Zoom meeting function, as opposed to the webinar function for the drop-in will allow the potential for breakout rooms and open text chat. Finally, it is worth exploring how registration will function for these sessions to allow for more visibility and for people to be able to register while the session is occurring.

Additional Partners/Portfolios

Bringing in additional partners and portfolios that play an active role in how research data is managed, but are traditionally not included in RDM training, would be a way to make the series more relevant and valuable to researchers, and also work to create a more holistic RDM support system at UBC. Potential partners could include Research Ethics, Indigenous Research Support Initiative, University-Industry Liaison Office, as well as others deemed relevant in future discussions.

Appendix A: Communications Channels & Materials

Appendix B: Course Descriptions

Session 1: Introduction to Research Data Management

Session length: 1.5 hours
Presenters: Susan Atkey
Helpers: Doug Brigham + ARC Specialists
In person delivery: Monday, September 28, 10:00am
Description:
This workshop will introduce the main concepts of RDM, including the research data life cycle, Tri-Council requirements, as well as supports and services offered at UBC. The course will then introduce learners to a mock project and dataset arranged in a file hierarchy with poorly named folders and files, and go through a hands on exercise to clean the structure and file names following best practices.

Susan Atkey is a Humanities and Social Sciences Librarian at UBC Library, working with Linguistics, Anthropology, Hispanic Studies, and Central & Eastern European Studies. Her Research Data Management focus is on the management of language and linguistics research data. She is a member of the Research Data Alliance (RDA) Linguistics Data Interest Group.

Learning objectives & outcomes:
Objectives
- Intro to RDM and data lifecycle
- Tri-Council Requirements
- Support at UBC
- Intro to mock project and data used through Fall Sessions
- Exercise to arrange file hierarchy with poorly named folders and files

Outcomes
- Identify stages of research data life cycle
- Identify Tri-Council requirements
- Identify RDM services and support at UBC
- Identify and apply best practices in file naming conventions and folder hierarchy

Session 2: First Steps in Data Management Planning

Session length: 1.5 hours
Presenters: Sheryl Adam
Helpers: Doug Brigham + ARC Specialists
In person delivery: Tuesday, September 29, 10:00am

Description:
Having a data management plan (DMP) is a great way to ensure that your research project remains on track and that your data retains its integrity throughout the entire research life cycle. This course will introduce the concept of a data management plan, and using the mock project data from session 1, will examine a 'dirty' dataset and explore what is needed to have high-quality descriptions and documentation of data, articulated in a CSV or readme file.

Sheryl Adam is the subject liaison librarian for Geography, Sociology and Psychology at Koerner Library. She has been presenting workshops on Research Data Management for more than four years.

Learning objectives & outcomes:
Objectives:
- Identify elements of high-quality variable level metadata
- Identify and articulate the value of a readme file

Outcomes:
- Clean up variable level metadata in a dirty dataset
- Develop a readme file for a dataset

Session 3: Introduction to the DMP Assistant

Session length: 1.5 hours
Presenters: Mayu Ishida
Helpers: Doug Brigham + ARC Specialists
In person delivery: Wednesday, September 30, 10:00am
Description:
The Data Management Planning (DMP) Assistant is a tool created for Canadian researchers to develop a data management plan, by asking a series of questions about what will be done with research data throughout its entire life cycle. Building off the work done in sessions 1 and 2, this course will introduce the DMP Assistant and walk through the mapping of its questions.

Mayu Ishida is a Science Librarian at Woodward Library, UBC Vancouver. She provides research support to students and faculty in biology, computer science, mathematics, and statistics, and teaches research data management workshops to graduate students.

Learning objectives & outcomes:
Objectives:
- Introduce the DMP Assistant and its purpose
- Introduce the questions asked by the DMP Assistant and where to go for support

Outcomes:
- Map mock project onto DMP elements
- Identify considerations for various DMP elements
- Identify supports and services when completing your own DMP

Session 4: Re-Introduction to the Data Life Cycle

Session length: 1.5 hours

Presenters: Allan Cho
Helpers: Doug Brigham + ARC Specialists
In person delivery: Thursday, October 1, 10:00am
Description:
This course will begin by discussing the DMP work from the previous day, encouraging conversation about differences across disciplinary data. It will then move on and discuss details of the data life cycle, the FAIR principles of data (Findable, Accessible, Interoperable, Reusable), and introduce Git and OSF as helpful data management tools. The session will conclude by preparing everyone for the next session, which will include having an OSF and Git account.

Allan Cho is the Research Commons Librarian. He coordinates a team of graduate students at the Research Commons and oversees the peer-led instruction and consultation spaces and services. In collaboration with Library and campus partners, Allan organizes the Graduate Student Workshop Series and programming designed to assist students with the research process.

Learning objectives & outcomes:
Objectives:
- Create OSF and Git accounts.
- Identify the data life cycle in detail within the context of the DMP
- Identify the guidelines of FAIR for digital assets

Outcomes:
- Describe the various stages of the life cycle:
    - Plan, Collect, Assure, Describe, Preserve, Discover, Integrate, Analyze
    - Use Cheese data as a case study in the 8 stages of the data lifecycle
    - Practice identifying elements of data lifecycle and their function
- Apply the FAIR principles in the context of the Fall Sessions data

Session 5: RDM Fall Series 2020 Office Hours # 1 (4 concurrent sessions for each of the week's workshops)

Session length: 2 hours
Presenters: Fall Series Team
Helpers: Doug Brigham + ARC Specialists
In person delivery: Friday, October 2, 10:00am - 12:00pm
Description:
Open office hours for any questions relating to this week's RDM Fall Series sessions. We will have a number of people available for you to talk to. So bring your questions!

Session 6: Introduction to Open Science Framework

Session length: 1.5 hours
Presenters: Mathew Vis-Dunbar & Susan Paterson
Helpers: Doug Brigham + ARC Specialists
In person delivery: Monday, October 5, 10:00am
Description:

RDM touches every aspect of the research life cycle. However, working across multiple platforms makes discrete connections across the research life cycle difficult. OSF allows us to draw these linkages and identify where each facet of the research life cycle and your data interact. In this session we will see how OSF addresses the issues of discrete connections across a project and have a hands on exercise to get you set up with a project page to start mapping out your research or your lab.

Mathew Vis-Dunbar is the Southern Medical Program Librarian at UBC Okanagan. As a member of the Fostering Open Science @ UBC initiative, he works with students, researchers, and colleagues across UBC in support of open, reproducible research practices.

Susan Paterson is a Humanities and Social Sciences Librarian in Koerner Library. Her liaison responsibilities include Social Work, French Language and Literature and Government Information. She is a member of the Library's Research Data Management team.

Learning objectives & outcomes:
Objectives:

- Identify how research data interacts with all aspects of the research data life cycle
- Identify how OSF connects otherwise disconnected aspects of the research life cycle
- Explore an OSF project with the structure required to connect one's data with each aspect of the research life cycle

Outcomes:

- Articulate the evolving role of data in the research life cycle and the unique considerations for management at each of these stages
- Describe the need to be able to map the discrete components of the research life cycle to best plan for RDM
- Build an OSF project that addresses the data organization and storage needs as data moves from raw to processed to ready for archiving

Session 7: Introduction to Data Repositories

Session length: 1.5 hours
Presenters: Marjorie Mitchell
Helpers: Doug Brigham + ARC Specialists
In person delivery: Tuesday, October 6, 10:00am
Description:
Data Repositories exist as trustworthy storage and access platforms for data arising out of research activities. Some are discipline specific while others accept datasets from the gamut of disciplines. In this session we will see samples of Canadian repositories, look at their specific uses, learn why metadata is so important, learn where to find metadata schema for your specific discipline. The session will also look at the common types of data repositories, including: Subject Specific and Structured Repositories; Institutional Data Repositories; and Unstructured or General Repositories. You will be able to locate and view existing datasets in the planning phases of your own research and better prepare to manage the data assigned to you. By managing your data with deposit and preservation in mind, you will apply the data management skills you have learned this far and set yourself up to meet funder, publisher, and research lab requirements.

Marjorie Mitchell is the Copyright, Scholarly Communications, and Research Data Management Librarian at UBC Okanagan. She has been presenting on Research Data Management topics since 2015.

Learning objectives & outcomes:
Objectives
- Re-intro data life cycle
- Re-intro Tri-Council requirement for deposits
- Intro different types of repositories and explain their differences / use cases
- Intro Registry of Research Data Repositories as a tool to find disciplinary repositories
- Intro TRUST principles of digital repositories

Outcomes
- Identify appropriate repository for your project's / session's dataset
- Identify value in depositing data and how it contributes to the scholarly dialogue
- Identify metadata as a critical element to making data discoverable
- Describe the role of repositories in the research data life cycle

Session 8: Dataverse as an Institutional Repository

Session length: 1.5 hours
Presenters: Doug Brigham
Helpers: ARC Specialists
In person delivery: Wednesday, October 7, 10:00am
Description:
How and why do you share your data with other researchers in your field, once your active research work is done? Maybe you've submitted an article for publication and the reviewers want access to your data. Maybe you want to share the data that supports your thesis or dissertation. Maybe you want to ensure that other researchers (yourself included) can access and use your data in the future. In this session, we'll look at UBC's institutional data repository, Dataverse. We'll examine some of the workflows that support depositing data, including data preparation, descriptive metadata and persistent identifiers.

Doug Brigham has been the Research Data Management Librarian at UBC Vancouver since 2020. He is the administrator for UBC's space within Dataverse@Scholars Portal and provides support to researchers in many disciplines. He is a member of the Research Data Alliance.

Learning objectives & outcomes:
Objectives
- Understand how institutional data repositories support researchers.
- Examine the work necessary to prepare data for deposit.
- Understand the role of descriptive metadata in making data accessible.
- Awareness of the support resources available at UBC.

Outcomes
- Prepare a dataset for deposit in Dataverse.
- Deposit and publish a sample dataset in Dataverse.
- Have a personal account on Dataverse @ ScholarsPortal, ready for your own data.

Session 9: Data Discovery and Deposit using FRDR, Geodisy and Globus

Session length: 1.5 hours
Presenters: Nick Rochlin
Helpers: Doug Brigham + ARC Specialists
In person delivery: Thursday, October 8, 10:00am
Description:
The Federated Research Data Repository (FRDR) is Canadian national data repository aimed to address a longstanding gap in Canada's research infrastructure by providing a single platform from which research data can be ingested, curated, preserved, discovered, cited, and shared. Geodisy provides map search functionality to supplement the FRDR discovery service, allowing researchers to find data based on location. This session will introduce data searching using the search bar as well as Geodisy, and will walk you through the steps of depositing data into FRDR using Globus File Transfer.

Nick Rochlin is the Research Data Management Specialist in UBC's Advanced Research Computing team. He is active in the Portage Network of RDM professionals, co-chairing the Training Expert Group and the Institutional Strategies Working Group, and an active member of the FRDR User Experience & Training Group.

Learning objectives & outcomes:
Objectives

- Outline FRDR's role in Canadian data repositories
- Outline use cases for depositing data into FRDR
- Introduce Globus as a file transfer solution
- Search for data using FRDR / Geodisy and play with search facets
- Assess quality of datasets found in FRDR

Outcomes

- Identify FRDR as a tool to search across Canadian data repositories
- Identify FRDR as a repository suited for data deposits larger than 5GB / 300 files
- Identify Geodisy as a tool for map-based data searches

Session 10: RDM Fall Series 2020 Office Hours # 2 (3 concurrent sessions for workshops 7-9)

Session length: 2 hours TBD
Presenters: Fall Series Team
Helpers: ARC Specialists
In person delivery: Friday, October 9, 10:00am
Description:
Open office hours for any questions relating to this week's RDM Fall Series sessions. We will have a number of people available for you to talk to. So bring your questions!

Session 11: Introduction to Sensitive Data & De-identification

Session length: 1.5 hours
Presenters: Nick Rochlin & Luc Letarte
Helpers: Doug Brigham + ARC Specialists
In person delivery: Tuesday, October 13, 10:00am
Description:
Data regarding human subjects can pose a number of unique challenges, and language that is articulated during the Ethics application and in consent forms has the potential to unnecessarily restrict what can be done with the data. This course will introduce considerations when dealing with sensitive data, including levels of access, consent forms, direct and indirect identifiers, and will introduce the process of data de-identification and considerations.

Nick Rochlin is the Research Data Management Specialist in UBC's Advanced Research Computing team. He is active in the Portage Network of RDM professionals, co-chairing the Training Expert Group and the Institutional Strategies Working Group, and an active member of the FRDR User Experience & Training Group.

Luc Letarte is a Security Analyst with the Sensitive Research team, at UBC Advanced Research Computing. He is a graduate from the University of Québec in Montréal and holds many years of experience in different segments of the public and private sector including financial, legal and health care. He specializes in information technology governance, privacy and security.

Learning objectives & outcomes:
Objectives
- Introduce considerations when dealing with sensitive data
- Introduce security and privacy considerations
- Discuss de-identification and potential issues with process
- Discuss considerations where depositing sensitive data into a repository

Outcomes
- Identify privacy and security considerations throughout data lifecycle
- Explore the language of consent forms in regards to data sharing
- Identify direct and indirect identifiers
- Identify tools/techniques for de-identification and considerations

Session 12: Introduction to REDCap

Session length: 2 hours
Presenters: Michael Tang
Helpers: Doug Brigham, Luc Letarte, Hannah Park
In person delivery: Wednesday, October 14, 10:00am
Description:
The REDCap (Research Electronic Data Capture) platform is a secure web application for building and managing research data collection instruments. The platform is specifically designed to support online or offline data capture for research studies. The UBC ARC REDCap platform includes two instances of REDCap are available to the UBC research community. We'll cover the REDCap basic features that you may need to meet your research projects' data collection requirements, that includes: building data instruments and surveys, preparing data set

exports, and managing user access control. We'll also share some creative ways REDCap can be set up to help facilitate your research data collection workflow.

Michael Tang is a Scientific Analyst with the Advanced Research Computing team at UBC. He has a Bachelor of Applied Science in Computer Engineering from UBC. Michael has over five years of experience administering and supporting researchers with the use of the REDCap platform.

Learning objectives & outcomes:
Outcomes

- Set up a research project in REDCap
- Build data collection instruments
- Create surveys for research participants
- Form datasets for export and analysis
- Facilitate electronic consent

Session 13: Data Horror Stories & Open Panel with UBC Researchers

Session length: 90 mins
Presenters: Nick Rochlin (host), Fiona McDonald, Jeffrey LeDue, Jason Pither
Helpers: Doug Brigham + ARC Specialists
In person delivery: Thursday, October 15, 10:00am
Description:
This session will be divided into two parts. The first part will feature UBC researchers discussing their experiences with data management, with a focus on unforeseen events that lead to a disruption, near disruption, or horror story in their research. The second part will feature a panel discussion in which UBC researchers will answer questions relating to data and RDM.

Dr. Fiona P. McDonald is an Assistant Professor of Visual Anthropology at the University of British Columbia, Okanagan where she is the co-Director of the Collaborative + Experimental Ethnography Lab (www.ce2lab.org) where her research focuses on Climate Justice and Social Change. Her current research focuses on water rights, cold climate housing, sensory ethnography, and open access digital publishing. She is a founding member of the Ethnographic Terminalia Collective, which curated exhibitions at the intersections of art and anthropology from 2009–2019. Her ongoing research looks at historical and contemporary uses of woolen blankets when transformed by contemporary artists, craft makers, and Indigenous communities in North American and Aotearoa New Zealand. She supervises undergraduate and graduate students in community-based research projects.

Jeffrey LeDue is the Managing Director of the NeuroImaging and NeuroComputation Centre and Coordinator of the Dynamic Brain Circuits Research Excellence Cluster at the Djavad Mowafaghian Centre for Brain Health at UBC Vancouver Campus. His expertise spans microscopy, imaging, optogenetics, custom hardware and data analysis pipelines. He has co-authored over 30 papers with 9 PI's since joining UBC in 2010. Prior to UBC, he was the Vision Science Imaging Specialist in the Department of Molecular and Cell Biology and School of Optometry at the University of California, Berkeley and worked on near-field optics and

biological atomic force microscopy in the Department of Physics at McGill University in Montreal.

Dr. Jason Pither is an ecologist and principal investigator in the Biodiversity and Landscape Ecology Research Facility at UBC's Okanagan campus. Since 2016, Dr. Pither has been collaborating with members of both UBC campuses to build capacity around training and support of best practices in Open Science. He is co-lead on UBC's strategic Fostering Open Science @ UBC initiative. At both campuses he has delivered more than 20 seminars, guest lectures, and workshops on Open Science and reproducible workflows using R and R Markdown. Nevertheless, has some of his own data horror stories to share.

Session 14: Office Hours (1 session for workshop 11)

Session length: 2 hours
Presenters: Fall Sessions Team
Helpers: ARC Specialists
In person delivery: Friday, October 16, 10:00am - 12:00pm
Description:
Open office hours for any questions relating to this week's RDM Fall Series sessions. We will have a number of people available for you to talk to. So bring your questions!