

The SSHRC DMP Workshop

A Final Report

Chuck Humphrey, Director of the Portage Network
March 2017

Table of Contents

EXECUTIVE SUMMARY	3
1.0 INTRODUCTION.....	6
2.0 OBJECTIVES.....	6
3.0 THE WORKSHOP.....	7
3.1 SCOPE.....	7
3.2 METHODOLOGY.....	7
3.3 PARTICIPANTS.....	9
4.0 OUTCOMES	11
4.1 DMPs.....	11
4.2 RESEARCHER QUESTIONNAIRE.....	12
4.3 LIBRARIAN QUESTIONNAIRE.....	14
4.4 CAPSTONE EVENT	15
4.5 CAPSTONE EVALUATION.....	17
5.0 ANALYSIS.....	18
5.1 THE SAMPLE	18
5.2 ASSISTANCE IN COMPLETING A DMP	20
5.3 DELIVERY OF THE CAPSTONE EVENT.....	22
5.4 ISSUES AND CHALLENGES IN COMPLETING A DMP	23
6.0 CONCLUSION	27
6.1 LESSONS LEARNED	27
6.2 LIMITATIONS OF THIS WORKSHOP	28
6.3 NEXT DEVELOPMENTS IN DMP SUPPORT.....	29
Appendix I: Agenda for the Orientation Video Conference	30
Contacts and Participants.....	31
Appendix II: Tasks and Deliverables	33
Appendix III: Agenda for the Capstone Event.....	34
Capstone Participants / Participant(e)s.....	36
Appendix IV: A Summary of the Analysis of Submitted DMPs.....	38
Appendix V: SSHRC DMP Workshop Survey	41
Using the DMP Assistant and Interface	41
Please rate your level of agreement with the following statements.....	41

Using the Template	41
Was the user interface of the DMP Assistant intuitive? Please note any issues you encountered.....	41
The length and coverage of the template is what I would expect when documenting a data management plan	41
Do you feel that using DMP Assistant prompted you to consider aspects of data management that you would not have included by yourself?	42
The questions within the sections are meaningful to my research.....	42
The questions are comprehensive and don't leave out any data management factors important to my research.	42
The terms, concepts and language used reflect those used in my research area or discipline.	42
Were there any questions that you could not answer?	42
Using the Guidance	42
The guidance text helped me to understand and answer the questions.	42
Would more specific guidance be helpful for any of the sections or questions?.....	43
Assistance	43
In completing your DMP, did you work with the local librarian identified for you?	43
Data Management Plans in Policy and Use	43
Please rate your level of agreement with the following statements.....	43
Appendix VI: SSHRC DMP Workshop Librarian Questionnaire.....	45
Appendix VII: Capstone Evaluation by Researchers and Observers	49
Appendix VIII: Notes of Audience Discussions at the Capstone Event	51

EXECUTIVE SUMMARY

The following report describes the results of a multi-phase Data Management Plan (DMP) Workshop led by Portage and the Social Sciences and Humanities Research Council (SSHRC) to explore the use of DMPs by a small group of SSHRC-funded researchers, to identify related data management issues and barriers in their community, and to assess the level of preparedness in their community to adopt DMPs on a wider scale. The findings of this Workshop are intended to inform SSHRC's data management policy development and implementation timelines.

Methodology:

In Fall 2016, SSHRC recruited Principal Investigators from eighteen research projects funded through the Partnership Grant and Insight Grant funding opportunities to participate in the Workshop. Selection of participants targeted disciplinary, regional and institutional diversity.

The Workshop had three phases:

- i. Invite participants to complete DMPs via the Portage network's *DMP Assistant*, with support available;
- ii. Collect issues and challenges through a questionnaire to participants; and
- iii. Convene researchers at a capstone event to solicit detailed feedback from participants and to promote a broad discussion about related challenges.

Findings and Outcomes:

DMPs and Questionnaire Responses

Analyses of the DMPs submitted by Workshop participants showed that researchers, with some support and direction, can successfully complete DMPs. Comments received from researchers also revealed many benefits from the completion of the exercise. Researchers found themselves being introduced to topics that they had not previously considered in managing their data. While certain results revealed among some researchers an incomplete understanding of a particular data management topic, responses also showed a new awareness of campus services at their institutions that could provide support.

An analysis of responses received for the researcher questionnaire, which was distributed to participants after the completion of the DMP exercise, revealed important findings about their experiences. Many of the researchers entered the exercise with unclear expectations of the length and coverage of a DMP and were unsure whether the set of questions provided was

comprehensive for their field of research. It is reasonable when preparing a DMP for the first time to be unsure about these items, as both require some form of benchmark. However, respondents were unanimous that the guidance text helped them to understand and answer the questions.

Responses to three items are particularly reassuring about the meaningfulness of the content covered:

- i. There were aspects of data management covered in the DMP template that the researcher would not have otherwise considered;
- ii. The content in the template was applicable to their research; and
- iii. There were no significant issues between the language in the template and their research area.

Library Directors at the participants' universities were approached to provide local support for participants, and all agreed to provide a local contact to work with researchers in completing their DMPs. This unanimous response from the academic library community reflects their commitment to support research data management at their institutions. Eleven of fourteen funded projects worked with the local librarian. The researchers and librarians all reported positive experiences.

Overall, these findings show the particular importance of the guidance text provided in DMP Assistant and a local contact in the library. A strategy to provide direct support goes a long way in helping researchers prepare DMPs.

Capstone Event

The in-person capstone event attracted the participation of twelve researchers and five data managers. Ten observers who represented research stakeholder organizations were also part of the discussions. The agenda for the event was designed to gather more detailed accounts of researcher experiences through small group discussions.

At the capstone event, researchers raised a variety of issues and challenges they faced as barriers to completing DMPs, which can be broadly summarized in three categories:

- i. **Resources:** Time, funding and technological needs are the primary impediments to researchers completing DMPs.

- ii. **Skills:** Researchers are being asked to address data management topics that stretch their training or research experiences.
- iii. **Culture:** Not all disciplines have an openly articulated data culture and as a result, researchers may adhere to a mix of norms and beliefs about data management that can impact their adoption of data management practices.

Despite these challenges, comments from researchers at the capstone event illustrated an openness to tri-agency data management policy, and reinforced the need for such a policy. Participants articulated that although they may have approached the DMP initially as an administrative task, they found the exercise led to improvements in their research plans and methodologies – completing a DMP benefitted their research. They also believed that DMPs would be useful in preparing for ethics approval.

Researchers highlighted the importance of a tri-agency policy with DM requirements as a key step in advancing the culture of data management in Canada, and recognized that data management is becoming an international best practice – several participants who are members of international teams noted that their collaborators are subject to requirements for DMPs. Other stakeholders in attendance noted the importance of a tri-agency policy with mandatory requirements as providing a rationale for institutions to invest in data management services and supports for researchers.

Concluding Thoughts and Next Steps:

While the Workshop showed the importance of providing researchers with assistance in completing DMPs, it also illustrated the need for local campus support organizations to work collaboratively in providing services to researchers and for raising awareness of the importance and methods of data management in general. The Workshop raised awareness of DMPs, related best practices for managing research data, and data management resources available at institutional libraries that can support and guide researchers in completing DMPs. It also revealed the benefits and obstacles researchers share with one another in managing and sharing their research data.

Through organizations such as Portage, Canada is part of a growing international community collaborating in the development of infrastructure to support DMPs. As DMPs continue to evolve internationally, their uses will expand, providing greater benefits to more stakeholders in the research community. SSHRC should continue to engage researchers, the Portage Network and other stakeholders as it continues to monitor and respond to these developments.

1.0 INTRODUCTION

Governments and research funders across the globe are becoming increasingly aware of the value of digital research data, the importance of fostering reuse of digital research data, and the need for policies to enable excellence in data stewardship. Canada has joined many other countries at the forefront of this movement, as shown in its support for the Organisation for Economic Co-operation and Development's [Declaration on Access to Research Data from Public Funding](#) (2004); its commitment to the [Open Government Declaration](#) (2011); and its approval of the [G8 Science Ministers Statement](#) (2013). In 2016, the Tri-Agency published a statement of principles to promote excellence in digital data management practices and data stewardship in agency-funded research.

SSHRC is promoting better data management and stewardship practices among its community of researchers and expects data collected through agency funds to be managed appropriately. The Council also recognizes that data management plans (DMP) are considered an international best practice and is exploring a data management policy that would require award holders to develop suitable DMPs. The aim of this project was to pilot the adoption of DMPs with a number of SSHRC researchers to identify issues and barriers in the community and to assess the level of preparedness of the community to adopt DMPs on a wider scale.

2.0 OBJECTIVES

The Workshop assembled a group of grant holders to better understand data management practices, supports, and challenges among humanities and social science researchers. The findings of this work will inform SSHRC's data management policy development and implementation timelines. In particular, this workshop had the following objectives:

- **Assist Workshop Participants in Completing DMPs** - Although developing a DMP in the initial stages of a research project has become an international best practice in the management of research data, knowledge of this practice varies across humanities and social science disciplines. Some researchers understand the elements of data management and develop DMPs as an integral stage of the research process, whereas other researchers are only beginning to learn the importance of data management, if at all. Moreover, support for managing data differs from institution to institution. Participants in the SSHRC DMP Workshop were asked to complete DMPs with support made available to answer their questions and assist them in completing their plans.
- **Document Issues and Challenges that Arise in Completing DMPs** - Given the diverse levels of awareness of data management among humanities and social science researchers and differences in the degree and type of support provided by research

institutions, various issues were anticipated to arise when participants completed the DMPs. This could range from knowledge gaps that pose challenges to completing the plans to concerns over lack of institutional capacity to store data being created during the research project. SSHRC sought to document the issues and challenges encountered during the Workshop. As such, participants would be expected to complete a short questionnaire about the experience of completing the DMPs and the challenges they would expect to face in managing their data.

- **Design and Conduct In-Person Capstone Event with Workshop Participants** - The first two stages laid the foundation for an in-person workshop with participants. This event was structured for participants to discuss their experience of completing DMPs and the information provided in the questionnaire. The participants were also to receive feedback about their DMPs. SSHRC required a design for the workshop that ensured fruitful discussion and feedback to the participants and effective facilitation.

3.0 THE WORKSHOP

3.1 SCOPE

The project consisted of 1) supporting participants in completing data management plans via the Portage Network's *DMP Assistant*, 2) identifying issues and challenges through a questionnaire to participants, 3) designing and conducting a workshop to provide feedback and discuss challenges, and 4) producing a report on Workshop findings.

3.2 METHODOLOGY

The planning and organization of the Workshop began in June 2016, while its delivery was undertaken over a three-month period running from October 31, 2016 to January 24, 2017. The steps conducted to complete this workshop consisted of the following activities:

- A. A sample of funded Insight and Partnership projects was identified by SSHRC and the principal investigators (PIs) holding these awards were invited to participate in the DMP Workshop. The PIs from 18 projects accepted this invitation. This phase was completed in October 2016.
- B. An orientation video conference was held with on October 31, 2016 to introduce the sample of researchers to the Workshop's objectives and DMPs. Participants were sent a copy of the slides from the presentation about the Tri-Agency Data Management Policy Initiative and about preparing DMPs using the Portage DMP Assistant. The participants were also sent copies of the Tri-Agency Statement of Principles on Digital Data Management and a summary of the feedback received on the Statement of Principles.

- C. Additional instructions about DMPs were sent on November 5, 2016 describing the local help that would be available to participants, reminding them of deadlines for each phase of the Workshop, walking them through a script in using DMP Assistant, and letting them know when the next communication would be sent.
- D. The next communication to the researchers was posted on November 14, 2016, covering background about developing a plan and describing the librarian service that had been arranged at their institution to provide them with local support in preparing a DMP.
- E. The library director at each participant's home institution was invited on November 3, 2016 to identify a local contact in the library to work with the researcher on her or his DMP. Upon receiving this information, a joint message was sent on November 16, 2016 to each researcher and her or his local library contact introducing them and offering any assistance they might like from the Portage workshop instructors.
- F. The Portage Research Intelligence Expert Group drafted a questionnaire between November 1, 2016 and November 28, 2016 for assessing the researchers' DMP experiences. SSHRC made some changes before settling on a final version of the questionnaire. The online questionnaire was administered by SSHRC, which was conducted starting December 23, 2016 and ran until January 16, 2017.
- G. A communication was sent on December 7, 2016 to the researchers with instructions about submitting their completed DMPs and answering the questionnaire assessing their DMP experiences.
- H. SSHRC sent participants instructions on December 14, 2016 about travel arrangements for the capstone event to be held on January 24, 2017.
- I. Messages were sent between January 3 and January 7, 2017 acknowledging receipt of submitted DMPs from researchers. Individual messages were sent to remind researchers who had yet to complete their DMPs or questionnaires.
- J. An analysis of the DMPs and questionnaires was conducted during the week of January 16, 2017.
- K. Between December 15, 2016 and January 13, 2107, the Portage instructors prepared a program for the capstone event that was reviewed and revised with feedback from SSHRC. The final version of the program and accompanying materials were sent to the researchers on January 18, 2017.
- L. The capstone event was held on January 24, 2017.

3.3 PARTICIPANTS

The frequencies in Table 1 are counts of funded projects, except the final row, which represents individual responses from participants who attended the capstone event. This included five responses from participants who were observers representing stakeholders in research data management.

Table 1: The Number of SSHRC Funded Projects by Workshop Activities and Type of Grant			
	Partnership Grant	Insight Grant	Total
Invited to participate	11	7	18
Completed a DMP	7	6	13
Completed Researcher Questionnaire	9	5	14
Completed Librarian Questionnaire	10	5	15
Completed Post-capstone Questionnaire[†]	9	3	17[‡]

[†]The frequencies for this activity are for participants at the capstone event and may include multiple researchers and data managers from projects in addition to observers attending the capstone.

[‡]This total includes five observers from stakeholder organizations.

SSHRC recruited PIs from eighteen funded Partnership and Insight grant holders for the DMP Workshop. Fourteen of these projects became the focus of this Workshop and provided the information about researcher and data manager experiences with DMPs. Of the original eighteen PIs, one did not engage in any of the activities. This researcher, however, did receive all communications sent to Workshop participants and was kept informed of the Workshop's progress. Two PIs withdrew in December 2016 saying that they did not have the time to dedicate to the Workshop. One PI was on sabbatical until the end of December 2016, indicating a possibility of catching up in January 2017 prior to the capstone event. While this researcher did not submit a completed DMP nor attend the capstone event, a completed researcher questionnaire was received, indicating that some work had been done in preparing a DMP.

In November 2016, Library Directors at the thirteen home institutions of the PIs were asked if they would like to participate in the DMP Workshop by providing a local contact in their library who might work with a researcher in completing a DMP. Three of these institutions hosted two projects in the Workshop. In total, sixteen projects ended up with a local library contact. One researcher was located in a U.S. institution and another researcher's

institutional affiliation did not have a library contact. Librarians were asked to respond to a questionnaire about their experiences in working with PIs on DMPs. Of these, fifteen responded for sixteen of the projects.

Table 2: Number of Projects by Discipline of Study and Phase of the Workshop			
Field of Study	Recruitment	Completed DMP	Present at Capstone
Anthropology	1	1	1
Archaeology	1	1	1
Communications/ Media Studies	1	1	0
Economics	1	1	1
Fine Arts	2	1	2
Geography	2	1	1
History	1	0	0
Linguistics	2	2	2
Literature	2	2	2
Political Science	1	1	0
Psychology	1	0	0
Social Work	1	1	1
Sociology	2	1	1
TOTAL	18	13	12

Thirteen different disciplines from the humanities and social sciences were represented among the eighteen projects recruited for the Workshop (see Table 2). Of the PIs completing DMPs, only History and Psychology were not represented. At the capstone event, the project in Communications and Media Studies was not present because of a prior team research commitment. Overall, this group of projects provided a good cross-section of both humanities and social science disciplines.

4.0 OUTCOMES

4.1 DMPs

An analysis of the content from the thirteen submitted DMPs was conducted to identify common themes in the answers provided and to gain a sense of the understanding and readiness to complete the questions in the Portage data stewardship template. A summary of this analysis was presented to the researchers at the capstone event to capture their feedback regarding these results and to confirm the interpretation of this analysis (see Appendix IV for this summary).

Overall, the researchers provided informative replies that demonstrated thoughtful consideration of the data management topics covered in the template. The feedback from the researcher questionnaire and the capstone discussions reinforces this observation. These findings show that DMPs can be completed by researchers, possibly with a little support. More importantly, however, comments from the researchers revealed that they directly benefited from completing a DMP. They found themselves being introduced to topics that they had not previously considered in managing their data.

Three particular outcomes were discovered in the analysis of this group's DMPs. First, the amount of detail given in their responses to questions varied across the template's seven sections. For example, the researchers gave thorough and lengthy replies to the section on Data Collection. This was clearly a topic about which they were well prepared to discuss.¹ Given the opportunity, researchers seem pleased to talk about their data. On the other hand, the last section about Ethics and Legal Compliance had the fewest responses and the least amount of discussion. Four of the projects did not see this section as applicable to their research. Others felt that their institution was responsible for questions about ethical or IP data-related issues. It was unclear if this was an assumption on the part of the researchers or if they had actually confirmed this. The researcher questionnaire following the DMP exercise provides some additional insight into the challenges that researchers faced in providing answers. The questionnaire results are discussed below in section 4.2.

In addition to an uneven level of detail provided across the sections of the template, some of the discussion revealed an incomplete understanding of a particular data management topic. For example, the researchers provided long lists of the variety of documents they had produced or planned to produce under the section on Documentation and Metadata. They were expected

¹ It should be noted that many of these projects had already collected their data. In these instances, the DMP exercise was a retrospective task. Rather than projecting what was planned to happen, they were describing what had been accomplished.

to identify how their data would be described, how they would document the context in which their data were being used, and what standards they would employ to package this information or metadata. The answers, however, tended to follow methods that typically are associated with knowledge mobilization and the dissemination of research findings. Some of these lists were from projects that had been completed or were near completion. Therefore, these researchers were reporting on the type of documentation they had produced rather than what they might produce. The distinction between information that describes their research findings and documentation or metadata about the data was missed in these instances. Retrospectively, using documentation that described their research outputs provided a fallback strategy. This is not in itself entirely wrong. Rather, it is an incomplete metadata record for their data. Furthermore, standards for describing data were overlooked in that context. In fact, only a couple of the projects mentioned actual standards employed in documenting their data.

Finally, the responses in the completed DMPs also revealed a new awareness of campus services at their institution that could provide them with support. This was particularly evident in the responses to the section on Preservation. This was a topic about which researchers were learning new possibilities. During the capstone, one researcher said that, for her, the whole DMP experience was a discovery about the preservation of her data.

4.2 RESEARCHER QUESTIONNAIRE

Researchers were asked to complete a questionnaire² about their DMP experiences after submitting their finalized DMP. Twelve researchers both submitted a DMP and answered the questionnaire; two of the remaining researchers completed a questionnaire but did not submit a DMP; and one final researcher submitted a DMP but did not complete the questionnaire. The result was a total of fourteen responses to the questionnaire.

The questionnaire was available to researchers in both official languages and addressed five topics about DMP experiences:

1. The use of DMP Assistant and its interface;
2. The use of the Portage data stewardship template;
3. The guidance text accompanying each question;
4. Assistance from a local library contact;
5. DMPs in policy and use.

The vast majority of the researchers agreed or strongly agreed with the seven statements about the usability of DMP Assistant and its interface. The two items about DMP Assistant's

² See Appendix V for a copy of the questionnaire and responses to the close-ended items.

collaboration features were not seen as applicable to several researchers. One researcher, who consistently responded negatively to these items, may have misinterpreted the direction of the scale. While this is speculation, the answers to this researcher's open-ended questions were all very positive.

The vast majority of the researchers found the six questions about the content in the template to have agreed with their experience. Two of the items had five respondents who were unsure that: (i) the length and coverage was what they expected of a DMP and (ii) the set of questions was comprehensive for their field of research. Both of these items require some form of benchmark to provide an answer. It is reasonable for someone preparing a DMP for the first-time to be unsure about these items. Responses to three items are particularly reassuring about the meaningfulness of the content covered in the data stewardship template: (i) there were aspects of data management covered in the template that the researcher would not have otherwise considered, (ii) the content in the template was applicable to their research, and (iii) there were no significant issues between the language in the template and their research area. The item asking if there were any questions that they could not answer showed the largest division among the researchers: eight said there were; five said there were not; and one was unsure. Four of the researchers with Partnership grants said that they were unsure how to answer some of the questions because not all of the details of their project were clear, yet. A few of the humanities researchers said that certain questions did not seem relevant to their research, e.g., the topics addressed in the Ethics and Legal Compliance section.

The respondents were unanimous that "the guidance text helped me to understand and answer the questions." They were divided whether further guidance incorporating more specific information from their institution or a funder would be helpful. They seemed somewhat to be supportive of more discipline guidance: seven were in favour; three against; and four unsure.

Ten of the respondents said that they worked with the local librarian who had been identified for them; four did not. A couple of the researchers already knew their local contact and had worked with her or him previously. For those who worked with their library contact, nine of them commented about this experience, all were positive:

- It was helpful to have access to an expert that could assist with finding proper answers to our DMP questions.
- The connection you provided was and will be very helpful moving forward. We already have one other meeting booked.
- She helped to describe in general terms what a DMP was, and to get me set up using the interface.

- The local librarian helped to clarify terminology as well as specific sections of the DMP. Discussion with the local librarian prompted us to also think of certain aspects of managing our data that we have not envisioned before.
- Helpful in planning how data will be stored and accessed, as well as privacy issues.
- Le bibliothécaire désigné nous a aidé dans un premier temps à mieux comprendre l'objectif de la démarche, à voir son caractère évolutif et à mieux comprendre certains termes spécifiques utilisés. Il nous a également aidé en faisant pour nous certaines recherches, notamment en lien avec les possibilités d'archivage des données.

Of the four who did not meet with their local contact, two said that they had ran out of time to schedule an appointment. One researcher was from a US institution, for which a library contact had not been identified. The fourth researcher said that she chose not to contact the library.

The researcher questionnaire concluded with four items about DMPs from a policy and usage perspective. All respondents agreed or strongly agreed that a DMP is helpful in planning a research project and that a platform like DMP Assistant with templates and guidance is useful. The item stating that a DMP would help prepare an ethics submission was supported by ten researchers but one disagreed and three thought it was not applicable. This finding aligns with the usage pattern found in the responses to the Ethics and Legal Compliance section mentioned in 4.1 above. The fourth item had the widest disagreement among the respondents. This was the statement that the researcher already had formal documentation that covered the topics in the DMP: two strongly disagreed and seven disagreed. Of the remaining five, four agreed with this statement, while one strongly agreed.

4.3 LIBRARIAN QUESTIONNAIRE

Following the submission deadline for DMPs, library contacts were asked to complete a questionnaire about their experience with a researcher from this Workshop. It is important to think of institutions in the context of librarian participation in the Workshop. The library contacts were identified according to the institution of the project PIs. Two of the eighteen projects fell outside the normal Canadian institutional context, leaving sixteen institutions. Three projects were from the same institution, resulting in a total of thirteen institutions in this Workshop. Librarian contacts were arranged for all thirteen and twelve of them completed the librarian questionnaire.

Nine of the twelve librarians met with a researcher and four of these nine meetings also included a data manager. The librarians initiated these meetings in five instances, while the researchers organized the meeting in the other four. Only one meeting was required in eight of these cases; the other case required two meetings. Eight of the nine librarians felt well to very

well about how these meetings went overall. One librarian was unsure. Similarly, eight of the twelve librarians were satisfied to very satisfied with the Workshop experience. Four librarians were unsure, but three of them had not met with a researcher or data manager. When asked if they would do anything differently the next time, eight of the twelve said 'no.' Appendix VI has a table of the questionnaire results and the answers given to open-ended questions.

4.4 CAPSTONE EVENT

Seventeen participants attended the capstone event from twelve SSHRC-funded projects: seven Partnership grants and five Insights grants. In addition, attendance at the capstone included seven observers from within SSHRC and ten from other research data management stakeholders: IDRC, NSERC, CIHR, CARL, Portage, CARA, CAREB, RDC, LCDI, and the Federation for the Humanities and Social Sciences. Finally, there were four speakers: two from SSHRC and two from Portage. In total, there were 38 participants, not counting the SSHRC program officers who dropped in to observe throughout the day.

The agenda was organized around four topics (see Appendix III). The first item of business presented the objectives for the day and included an introduction of the participants at the capstone. The second item provided a review of research data management trends followed by small-group discussions about any changes that the researchers are experiencing in data management. The third agenda item involved a review of an aggregate analysis of the researchers' completed DMPs. This was followed by small group discussions about researcher reactions to DMPs. The fourth topic focused on SSHRC data policy development and the use of DMPs as part of a policy. The participants were asked to discuss the responsibilities associated with data management policy and about where to go next in formulating such policies.

Prior to the capstone event, researchers were asked to send a sentence or two about any surprises they encountered while preparing their DMP. These statements were included on the slide used to introduce the researcher. The following six statements were received prior to the capstone.

- I liked the Portage Guidance that accompanies each question. I don't think people always understand why they are being asked to write a data management plan. A common mistake I have seen in NSF proposals is DMPs that focus too narrowly on how data access will be limited (for human subject protection) and don't give enough information about how data will be preserved and shared and made useful to others. I think this system can play an important role in educating researchers about all the different ways they need to think about their data management and guiding them as they plan their research.

- As we are at the beginning of our project, the DMP process really helped us to plan what we want to do with our data and how we want to proceed. We found it very useful. However lots of points will need to be clarified.
- The process of writing the DMP really pointed out the great diversity of the data we are dealing with within our project, and showcased the importance of having a distinct data management approach for each type of data, and the challenges that come with it.
- I came to see the DMP process as about planning for the preservation and archiving of data, not about its ideal presentation or optimal accessibility.
- What surprised me most about the process was the extent to which our group had already developed many of the DMP elements. The process of incorporating these existing elements into a single document, however, provided us the opportunity to better appreciate the points at which those elements can be brought together more effectively. The exercise was tremendously valuable in that way.
- Nothing surprised me.

During their introductions at the event, researchers who had not provided a statement made verbal comments about their DMP experience.

- DMPs have evolved out of the need to create open access to research outputs. Creating a DMP has greatly helped guide that process.
- The whole movement to share data and conserve it has not been part of his area of study, which complicated the DMP process.
- A steep learning curve exists around creating a DMP; and as a result, more processes have to be figured out than were actually covered by the DMP.
- DMPs were particularly helpful in identifying practices that they had not considered. For example, they had not discussed who would assume responsibility for the data if someone critical left the project. Also, they had not planned how knowledge would be retained through this process.
- Due to the diversity of partners in his research, they had to spend a lot of effort thinking how to combine all the collected data. The DMPs were very helpful, but he faced a challenge with how to integrate the quantitative, qualitative, and aboriginal traditional knowledge data.

- In writing her DMP, she realized that many of the processes, while already being done, were informal. She realized in the process of completing the DMP that, in her field, methodological publications were being used to retain some of this otherwise tacit knowledge.
- Her partnership project required organizing the efforts of nine colleagues to complete the DMP. At the end of the effort, she realized that the formalization of these various processes was very beneficial and almost a research project in and of itself.

Appendix VIII provides a set of notes summarizing the discussions held around the second through fourth agenda items.

4.5 CAPSTONE EVALUATION

All of the participants who attended the capstone event received a short online questionnaire asking them to evaluate the day (see Appendix VII). Twelve of the seventeen researchers and data managers and five of the ten observers from RDM stakeholder organizations provided responses. Four questions were asked about the event:

1. Did the capstone event meet the participant's expectations?
2. How important was the information at the event for the participant?
3. Was the participant satisfied with the opportunity to express her or his thoughts during the event?
4. Overall, how did the participant rate the event?

Half of the researchers felt that the capstone event had met all of their expectations while the other half said that it met some of their expectations. All of the invited observers reported having their expectations met or exceeded. Eleven of the twelve researchers saw the information about DMPs and data policies as important or very important. One researcher found this information not very important. Four of the observers felt the same as the majority of the researchers, while one had no opinion. Both the researchers and observers were satisfied or very satisfied with their opportunity to speak at the capstone. One researcher was only somewhat satisfied. Sixteen of the seventeen respondents rated the overall event as good or very good, while one researcher rated it as poor.

Respondents were invited to comment on their evaluation of the event, of which seven of the seventeen offered additional information (see the comments in Appendix VII). One researcher, who was the source of the only low scores on all four questions, offered an explanation for his scores. Paraphrasing, he felt the Workshop could have been accomplished using video

conferencing. He also questioned the decentralization of research data management and thought that libraries were inappropriate and that a national system makes more sense. In contrast, one of the other respondents felt that the Workshop was a clear demonstration “that the views of the researchers and the broader stakeholder community are important to guide the development of data management policy and practice.” She saw this as an excellent way of gathering valuable input from researchers and stakeholders. Another researcher felt that the small group discussions should have been managed better to address the questions that had been asked. However, another researcher said that he found working in the discussion groups on focused questions very valuable.

5.0 ANALYSIS

The following analysis addresses the three objectives of the Workshop described in Section 2.0. Specific indicators have been identified to assess the degree to which each objective was achieved. The output of the Workshop discussed in Section 4.0 and content from some of the appendices are used as input for these indicators.

5.1 THE SAMPLE

An analysis of the sample of SSHRC-funded projects in this Workshop is an important factor in setting a meaningful benchmark to assess the three Workshop objectives. The selection of PIs and projects was not based on a representative sample of all SSHRC funded research nor was there ever an intention to generalize the outcomes to all humanities and social science researchers. Rather, the focus was to provide a diversity of disciplines in the humanities and social sciences that would allow exploring a range of data management issues across these fields of study. Every discipline has a degree of explicit and implicit norms regarding data management. These usually are learned in graduate courses in research methods or while conducting research. A better understanding of such norms provides insight into data management practices and an openness to preparing DMPs. Therefore, it was important that this Workshop have a variety of humanities and social sciences disciplines to capture some sense of the differences in norms.

Four indicators were used to assess the degree of diversity across the activities of the Workshop. This includes the range of disciplines in the initial sample and the completion rate of workshop activities by researchers in these disciplines. Projects either were Partnership or Insight grants, which is another way in which diversity of humanities and social sciences research can be expressed. The completion rate of recipients in these two grant categories provided a second indicator. Third, there is the overall completion rate of participants across Workshop activities and the fourth indicator is the engagement of library directors in

identifying a local library contact for DMP support. This latter indicator shows the evenness of support provided across the participants in the sample.

The sample of SSHRC-funded projects in the Workshop provided a good cross-section of research areas in the humanities and social sciences. As shown in Table 2, the eighteen recruited projects were from thirteen disciplines. Completed DMPs were provided for projects in eleven of the thirteen disciplines. The two disciplines for which activities were not completed were in History and Psychology. The capstone was short one other discipline, although the researcher in this field had completed a DMP and questionnaire. The overall completion rate of activities across disciplines upheld the majority of the initial cross-section of humanities and social science research. Furthermore, participation by grant type was also maintained throughout workshop activities. Two of the three projects in which the PIs did not complete any of the activities were Partnership awards, but the overall number of Partnership grants was four more than Insight grants. Of the twelve researchers who attended the capstone event, five held Insight grants, while the other seven were recipients of Partnership grants. This still honoured the original split between grant categories.

Two other factors – both part of the SSHRC sample design – were present in the thirteen completed DMPs. The regional distribution of these projects was a factor: one project was located in Atlantic Canada, two in Quebec, four in Ontario, three in the Prairies, two in BC, and one in the U.S. The size of the institution with which the PIs were affiliated was the other sampling criterion. Of those who completed a DMP, six researchers were from large universities, five from medium-sized institutions, and two from smaller universities.

Participation of PIs across activities was high with only a couple of exceptions. One researcher never responded even though this PI was included in all Workshop communications. Two PIs informed us in mid-December 2016 that the time commitment was greater than they could make and therefore withdrew. A fourth researcher was on sabbatical through December 2016 but did respond to the questionnaire. Another PI began a sabbatical in January 2017 but submitted a completed DMP and questionnaire before starting his academic leave. Fifteen of the PIs contributed to one or more of the activities with twelve completing all of them.

Library Directors at thirteen of fifteen institutions in the Workshop were contacted. Three institutions had two projects each in the Workshop (thus fifteen institutions in total instead of the eighteen for projects); one PI was located at a U.S. institution; and the final researcher was at an institution without a library director. All thirteen of the Library Directors who were approached provided a local contact to work with the researcher or researchers at their institution. This unanimous response from the academic library community reflects their

commitment to support research data management at their institutions and is a valid indicator of the evenness of support provided to the PIs in the Workshop.

A comment was made during the capstone that there was likely a bias in this group of researchers toward openness to data. Because the projects were not drawn as a representative sample, bias was less of a concern to the Workshop organizers than about who was involved. Nevertheless, knowing any potential bias is always helpful. This self-awareness of a particular bias can guide the interpretation of the feedback provided by PIs.

5.2 ASSISTANCE IN COMPLETING A DMP

As described in Section 3.2 on the methodology of the Workshop design, a variety of assistance was provided to Workshop participants. First, an orientation video conference call was held on October 31, 2016 to introduce the participants to the objectives of the Workshop, to DMPs, and to using DMP Assistant. A series of instructional messages were sent in November and December 2016 and in January 2017. Local assistance was provided to the researchers through a library contact, starting on November 16, 2016. Finally, the two Portage instructors were also available for consultation and did respond to a few of inquiries from researchers or their data managers.

The rate of completed DMPs and the response rate to the researcher questionnaire serve as two measurements regarding the right level of assistance to researchers. With three withdrawals from the original eighteen projects, there were thirteen completed DMPs out of fifteen projects. There was evidence that some work had been done on the two remaining DMPs, even though they were not submitted. One of these two PIs completed the researcher questionnaire, while the other participated in the capstone event. Fourteen out of fifteen PIs responded to the researcher questionnaire.

The participants in the Workshop were unanimous in the researcher questionnaire about the usefulness of the guidance text in DMP Assistant. Having in-context guidance while working on their DMP was seen as very helpful. They were more evenly divided about having further guidance directed toward their institution or the funding agency, although more support was indicated for discipline guidance. This is valuable feedback for the Portage DMP Expert Group when looking to improve on the guidance text of their data stewardship template.

Ten of the fourteen researchers also spoke with a local library contact at least once. This was corroborated in the librarian survey. Most of this support was provided through a single meeting. Only one librarian reported meeting twice with a researcher or data manager. The

following comments were made about how their discussions with a librarian proved to be helpful.

We had a general discussion of the services that are offered by my home institution as well as the project that we will be undertaking.

Helpful in planning how data will be stored and accessed, as well as privacy issues.

The local librarian helped to clarify terminology as well as specific sections of the DMP. Discussion with the local librarian prompted us to also think of certain aspects of managing our data that we have not envisioned before.

I met with the local librarian identified to our project. We also had a meeting/training session about Portage. It was helpful to have access to an expert that could assist with finding proper answers to our DMP questions.

I had some contact with the librarian identified here- I already knew him and had spoken to him about my project before. I also attended a data management workshop that he ran for my faculty.

She helped to describe in general terms what a DMP was, and to get me set up using the interface.

Le bibliothécaire désigné nous a aidé dans un premier temps à mieux comprendre l'objectif de la démarche, à voir son caractère évolutif et à mieux comprendre certains termes spécifiques utilisés. Il nous a également aidé en faisant pour nous certaines recherches, notamment en lien avec les possibilités d'archivage des données.

New departments are presently being formed, that we did not necessarily have direct contact with. The connection you provided was and will be very helpful moving forward. We already have one other meeting booked.

The local librarian assisted with access to the system.

Two respondents commented that they had not met with a librarian because it was not applicable in their situation and two other researchers said that they had difficulty scheduling a time to meet. However, the overwhelming response was positive about working with a local librarian to complete a DMP.

A couple researchers mentioned that they already had a working relationship with a local librarian and in some instances it was the same contact who had been provided through the Workshop. Portage considers the local working relationship between a researcher and a librarian as vital for a successful implementation of a DMP requirement. Toward this end, Portage strives to encourage researchers and librarians to work together on data management planning. If Portage experts are invited to assist a researcher directly, help will be provided but

only as part of a team consisting of the researcher, a local librarian, and a Portage expert. The focus is to strengthen the local relationship.

The indicators used to assess the assistance with DMPs show the particular importance of the guidance text in DMP Assistant and a local contact in the library. A strategy of having support in context that is available on demand, such as guidance text, and someone local who can help describe the process and interpret the language used in data management goes a long way in helping researchers prepare DMPs.

5.3 DELIVERY OF THE CAPSTONE EVENT

The purpose of making the capstone an in person event was to gather more detailed accounts of researcher experiences through small group discussions. While the event was also an opportunity to provide researchers with feedback about their DMPs and questionnaire responses, the primary focus was to learn through small group dialogue. Participation in the capstone consisted of seventeen participants from seven Partnership and five Insight grants (see Section 4.4). Five data managers from Partnership grants also participated in the capstone.

Ten observers who represented research stakeholder organizations were also part of these discussions. These participants contributed a sense of the wider community in which research occurs. Local campuses support researchers through a number of professionals who assist with administering research applications and funding, who contribute to necessary research processes (e.g., ethics approval), who provide research resources (e.g., research literature, data, or computing), and who bring expertise to the research process (e.g., data management or technical skills). Many of the questions in the data stewardship template suggest to researchers possible local campus partners who can assist in the management of their research data.

Five observers represented professionals on local campuses, such as, researchers, ethics officers, research service officers, librarians, and research data management specialists. These participants were from the Federation for the Humanities and Social Sciences, the Canadian Association of Research Ethics Boards, the Canadian Association of Research Administrators, the Canadian Association of Research Libraries, and the Portage Network, respectively. An invitation was also extended to the Canadian University Council of Chief Information Officers but regrets were sent just prior to the event. Two national organizations associated with research data interests were also observers: the Leadership Council on Digital Infrastructure and Research Data Canada. Finally, three other granting councils participated: NSERC, CIHR, and IDRC.

The design of the capstone event (see Appendix III) was to stimulate discussion around three topics.

- How are trends in research data management changing the way in which research is conducted in your field? How might DMPs help you adjust to such changes?
- From your experience in completing a DMP, what are the best ways that researchers and funding agencies can make use of them?
- Who will assume responsibilities for data management as it becomes expressed in policies and what role might DMPs play in this?

These discussions occurred at three tables of twelve to thirteen participants. A rapporteur was chosen at each table to summarize a group's discussion, which was shared with the full group. Summary statements from these discussions are reported in Appendix VIII.

A brief, online questionnaire was distributed after the capstone to all participants, asking them to help evaluate the event. Section 4.5 describes the outcomes of this evaluation (see Appendix VII for questionnaire results). The overall evaluation of the Workshop (question #4) showed eleven of twelve researcher respondents rating the event as good or very good, while all five of the observer respondents gave these same scores.

Another important evaluation criterion was whether participants felt they had been heard in their discussion groups. They were asked if they were satisfied with the opportunity to express their thoughts during the event. Fifteen of seventeen respondents were satisfied or very satisfied, while one was somewhat satisfied and one had no opinion. One respondent did express concern that the discussions drifted from the questions provided by the capstone organizers. She felt that the groups could have been managed better. This was not a criticism about having an opportunity to speak but rather reflected a level of frustration with some discussions not staying on topic.

5.4 ISSUES AND CHALLENGES IN COMPLETING A DMP

This section looks at a variety of challenges that researchers mentioned during the Workshop as barriers to completing their DMPs. These issues can be grouped into three categories: resources, skills, and culture. This section examines issues within each of these categories and provides evidence from the Workshop that offsets some of these concerns.

RESOURCES. It is said that the single, most valuable commodity to a researcher is time. This is a scarce resource that they strive to protect. As such, time was identified as an impediment to completing a DMP. The researchers of the two projects that withdrew from the Workshop both said that they did not have the time to complete the exercise. A lack of time, however, can be the product of competing priorities. One researcher was preparing for a sabbatical in the midst

of completing his DMP. He not only submitted his DMP on time, he completed the researcher questionnaire.

Some researchers see a DMP as a questionable administrative duty. As such, many researchers see this as a time sink. There is evidence from the U.S. that researchers may spend up to 40 percent of their time on administrative tasks related to their research. At this level, administration can be a burden on the researcher. One of the goals of a DMP is to share the information it contains with other administrative systems. The guiding principle is to enter the information once but reuse many times. For example, the data stewardship template contains several items that could be submitted in conjunction with an ethics application regarding the protection and treatment of a project's data. Developing an approach to the ethical management of the data in a DMP could serve multiple purposes. If the same information can be easily shared with different administrative requests, the burden of producing the information is reduced. Such information may also be more comprehensive as a result.

While the researchers were not asked to track the amount of time that it took to complete a DMP, the timetable for the Workshop allowed a six-week period to finish the assignment, giving them a fair amount of flexibility to fit this activity into their schedules. A few of the Partnership grant recipients had to coordinate their DMPs with other colleagues, which meant juggling schedules with others. This was mentioned as problematic in one instance.

Because the Workshop was confined to a three-month period, it was not possible to look at time expended on all data management activities -- including planning -- across the full life of a research project. In other words, planning for data management should be seen as part of the normal activities of managing data. In this context, the time spent on a DMP is time invested in the outcome of the project's data. It is hypothesized that a DMP will save a researcher time over the life of a project by preparing her or him in advance for activities involved in the transfer of data from a project to a repository for long term stewardship. If steps are not taken during a project that help prepare the data for deposit, the level of activity needed to accomplish this at the end of a project is vastly more time consuming.

Having adequate funds to hire a data manager was also mentioned as a resource impediment to data management. Some researchers were unaware that SSHRC allows expenses for data management. It was mentioned during a discussion, however, that an application receiving less funds than requested will often result in a researcher sacrificing data management to offset the shortfall. A DMP can help a researcher justify her or his request for data management support. It was mentioned that having examples of such costs would be helpful when putting together an application. Until more researchers are successful in receiving funds for data management

activities and can report accurate costs for them, applicants have difficulty knowing how much they should be requesting.

A slightly different resource challenge coming out of a couple of the humanities projects was a concern about the ongoing need for technology to support specific research outputs. For example, a humanities researcher had produced a searchable database and wanted to know how the search engine could be preserved along with the digital content in the database. Another researcher had a website that represented the output from his research. The researcher's data was used to display the website but he felt that without current web technology to display his results, the project's data would be incomplete. This is an interaction between form and content, where both are deemed necessary to be complete. While this complexity should be expressed in a DMP, it is not a barrier to preparing one. It seems to make the case for developing a DMP all the more significant. A humanities researcher in this situation needs to make the case that form and content together constitute the DMP.

SKILLS. Researchers are being asked in some instances to address data management topics that stretch their training or research experiences. How researchers respond varies. If they are in a team setting, someone might be included on the team with the skills that are missing. If they are an individual researcher, this challenge becomes a bigger obstacle. One option is to seek funding for a data manager. A researcher may choose to upgrade her or his data management skills (see the discussion about online training in section 6.2), but this would require a time commitment (see the discussion about time as a resource immediately above). There are also research data management contacts in local libraries who can provide assistance or who can make a referral to someone with whom the researcher could consult. A DMP does not require a definitive answer to all questions if one is not yet known. However, in those instances, a strategy that the researcher would follow to get an answer is an appropriate response.

CULTURE. In this context, culture refers to the norms and beliefs of accepted data management practices within a discipline. Not all disciplines have an openly articulated data culture and as a result, researchers may adhere to a mix of norms and beliefs about data management. Some of the cultural topics that occur most frequently include definitions of data, data management incentives, and privacy and ethics restrictions on data.

The issue about what constitutes data often occurs in discussions about data management. If data cannot be defined for a discipline, there is little justification for a DMP. One researcher wondered if working with a primary resource, such as an archival document, constitutes the creation of data. It was assumed that the act of creating data was necessary to qualify as research data. The definition for research data provided in the Statement of Principles on

Digital Data Management allows for a wide variety of evidence or artistic materials in research. This issue seems to be more contentious in discussions about data management than in what happens in actual practice. The highest quality of responses received in Workshop DMPs were the descriptions of the researchers' data. While their understanding of metadata was less evident, they had a clear grasp of what the data was in their projects.

What motivates researchers to share their data? This issue manifests itself in discussions about the ownership of data and incentives to share data. Increasingly, data ownership arguments are being countered by the need for multiple data stewards to ensure long-term access to data. A data owner can transfer stewardship rights to another party without relinquishing ownership. The more widely this norm becomes accepted the less the issue of ownership will remain a barrier. The motivation for a researcher to share her or his data then shifts away from discussions of ownership to one about incentives. Typical sentiment is that the academic rewards system is necessary to drive research data management forward. Researchers need fair recognition for the data they produce and share. Proper attribution given to the source of data is an ethical practice to be strongly promoted. But data sharing is occurring in some disciplines without being driven by the academic rewards system. In those disciplines, there is an underlying culture of data sharing, i.e., the norm is to share your data with others.

Another issue that is frequently raised as a barrier in data management, especially involving human participants in the research, is privacy and ethical constraints in the data. A common statement is that "the ethics board requires that I destroy my data". If consent does not allow for the possible reuse of the data nor have practices been proposed to protect human participants against disclosure in an ethics application, the result is largely predetermined that the data will need to be destroyed. A DMP increases the options to consider when planning for the management of confidential data. Continuous data monitoring in health research is revolutionizing not only data gathering techniques but also changing the practices for obtaining consent from participants. The old default of destroying confidential data is now being challenged by new technology that manages the risk of disclosure. A DMP can help a researcher plan how such risk-management technology can be implemented in their project.

One researcher raised a concern that SSHRC is attempting to push the humanities into a data-intensive research paradigm. A DMP template that favours data management topics more associated with the sciences or social sciences may contribute to such an interpretation. Nevertheless, the experiences in the Workshop revealed common ground among humanities and social science projects for many of the sections of the data stewardship template. The fourth point in Section 6.2 admits that work could be done to validate data management topics in the humanities further. Increasingly, humanities research involves organizing digital research

materials into collections. Nevertheless, these collections still require some form of management that should be planned prior to the research commencing.

The uneven level of detail in DMPs discussed in Section 4.1 points to the need for ongoing discussions at the local level while developing a DMP. A review process might be introduced to address this issue. The Workshop did not undertake a critical review of each DMP, providing each researcher with feedback about possible improvements. Regardless of how DMPs are implemented, an institution may offer a service that provides researchers with critical feedback on their DMP. This would permit researchers an opportunity to make adjustments to their DMP.

6.0 CONCLUSION

6.1 LESSONS LEARNED

The group of PIs and data managers who volunteered for this Workshop provided valuable insight into the challenges they faced in completing a DMP, giving SSHRC an opportunity to observe firsthand experiences from a variety of research areas in the humanities and social sciences. In particular, the Workshop demonstrated the value of a DMP to the researchers themselves. Many of the participants commented that they encountered data management topics that they would not have necessarily considered previously, although in hindsight, they see the value of addressing these topics in a DMP.

The Workshop also showed the importance of providing researchers with assistance through a local library contact. These contacts were further supported through Portage and in some instances, they consulted with Portage experts about the advice to provide. The inclusion of observers from the organizations representing local campus support for research was also important because of the emphasis it placed on the community needed by researchers to plan their data management activities. It also illustrated the need for local campus support organizations to work collaboratively in providing services to researchers.

The Workshop also revealed to the researchers, who often find themselves working in silos without much interaction with colleagues until they publish their findings, the common ground that they have with one another in managing their research data. They also saw data activities that were similar whether the research was in the humanities or the social sciences and they could speak with one another about optional practices in their context, as well as what might constitute a best practice. Getting researchers and data specialists engaged in conversations about best practices in managing data is a significant interaction for future developments in policies around data preservation and sharing. The desire exists to get these conversations

happening within disciplines in the humanities and the social sciences but it must also occur at the federation level of the societies representing these areas of study.

6.2 LIMITATIONS OF THIS WORKSHOP

There were limits to what could be learned about researcher adoption of DMPs in this Workshop and about how DMPs might be implemented as part of the granting process. Five of these limitations are discussed below.

- First, the Workshop did not address different incentives for researchers to complete DMPs. PIs from SSHRC-funded projects were asked to participate as volunteers. While these willing participants provided lessons about researcher experiences in completing a DMP and about the support they required, no systematic insights were gained into the incentives needed to engage researchers reluctant to complete a DMP.
- Second, the focus of the Workshop was around the experience of producing a DMP without a timeframe that permitted an examination of the longer term benefits of a DMP, such as, the time a DMP might save researchers during a project or the number of datasets deposited at the end of a project that are shared with others.
- Third, the Workshop did not explore different ways in which DMPs might be implemented in the research process. Questions remain about how DMPs should be incorporated into the grant application process. Should the researcher be required to submit a DMP? Should a submitted DMP be included in the adjudication process for a grant? Should the institution be responsible for administering a DMP, similar to the processes used for ethics approval? How frequently should DMPs be reviewed and updated? Who would be responsible for reviewing DMPs? What kind of mechanisms should be introduced to ensure that data are deposited at the end of a project in accordance with the terms expressed in a DMP?
- The Workshop only employed a generalized data stewardship template in DMP Assistant. This template reflects a universal set of best practices in data management across the data lifecycle, but did not focus on practices within specific domain areas. For example, there was no examination of specialized data management topics in the humanities or the social sciences.
- The researchers recognized the provision of a local contact in the library and the guidance text accompanying the data stewardship template to be valuable resources in completing their DMPs. However, alternative methods of providing researchers with support were not systematically explored in this Workshop. For example, would an online course about research data management have been helpful to researchers prior to beginning their DMP or to meeting with their local library contact?

6.3 NEXT DEVELOPMENTS IN DMP SUPPORT

Through Portage and Research Data Canada, our country is part of a growing international community collaborating in the development of infrastructure to support DMPs. Portage is contributing to a multi-national software development project with the Digital Curation Centre in the U.K. and the California Digital Library in the U.S., enhancing the online web application used by researchers to author DMPs. A meeting of this development team in February 2017 set priorities for enhancements to the common codebase for this application. The next major release of DMP Assistant will incorporate these new features, including a better editor for entering text in templates, a more robust method of sharing DMPs among applications (e.g., an API exchange with an RSO administrative application), a simplified data model to facilitate information exchanges with other systems, and publishing functionality enabling DMPs to be treated as a research output. Several European countries are beginning to use this DMP platform for their own national purposes. For example, Finland introduced mandatory DMPs in 2016 using the same platform that Portage supports. With this widespread adoption of the same platform, more co-development is beginning to take shape to advance best practices in DMPs.

The Research Data Alliance (RDA) has an Interest Group on Active Data Management Plans that is looking at greater integration of DMPs into data management activities in research. The software developers mentioned previously are also part of the discussion to create Active DMPs. The ideas that have been discussed to date look to combine project management techniques with research workflows to support data management in research. Research Data Canada is the official organization representing Canada at RDA and is our conduit to these discussions, although anyone can join an RDA Interest Group as an individual participant.

As DMPs continue to evolve internationally, their uses will expand, providing greater benefits to more stakeholders in the research community. This is an activity that SSHRC should monitor in conjunction with Portage and Research Data Canada.

Appendix I: Agenda for the Orientation Video Conference

Agenda

SSHRC Data Management Plan Workshop Orientation

Meeting Time: Monday, October 31, 1:00pm – 2:00pm (EST)

How to Connect: connecting to Video will allow you to view the presentation; connecting to audio will allow you to participate in the discussion.

Video

To join the meeting, click here: [1 SSHRC-CRSH DATA MANAGEMENT / GESTION DE DONNEES](#) or enter <https://sshrcnserc.megameeting.com/guest/?id=1026-556783> in your web browser.

Meeting Domain: sshrcnserc.megameeting.com

Meeting ID: 1026-556783

Audio

Local Dial-in number 613-960-7510

Toll-free Dial-in number 1 877-413-4781

Conference ID 1007107

Agenda Items

- **Welcome / Introductions (5 min)**
- **SSHRC's Data Management Policy Initiative (10 min)**
 - Jeremy Geelen, Policy Analyst, SSHRC
- **Portage's DMP Assistant (15 min)**
 - Chuck Humphrey, Director of Portage
- **Questions and Discussion (25 min)**
- **Next Steps (5 min)**

SSHRC DMP Workshop

Contacts and Participants

Contacts

Focus group activity (aims, timeline, etc.): Jeremy Geelen – Jeremy.geelen@sshrc.crsh.gc.ca

Workshop logistics (travel, etc.): Jacques Critchley – Jacques.Critchley@sshrc-crsh.gc.ca

Data management plans/questionnaire: Chuck Humphrey – humphrey@datalib.library.ualberta.ca (or institution rep, as appropriate)

Participants

Partnership Grant Holders

Name	Project	Discipline	Institution
Line Chamberland	Savoirs sur l'inclusion et l'exclusion des personnes LGBTQ (SAVIE-LGBTQ)	Psychology	UQAM
Ratana Chuenpagdee	Too big to ignore: global partnership for small-scale fisheries research	Geography/ Interdisciplinary	Memorial
Geoff Cunfer	Sustainable farm systems: long-term socio-ecological metabolism in western agriculture	History	Saskatchewan
Martin Goyette	Étude longitudinale sur le devenir des jeunes placés au Québec et en France	Sociology	École nationale d'administration publique
Timothy Harrison	Computational research on the ancient near east: an archaeological data integration, simulation, and 3-D visualization initiative	Archaeology	Toronto
Gary Libben	Words in the world	Linguistics	Brock
Julio Mercader Florin	Stone tools, diet, and sociality at the dawn of humanity	Anthropology	Calgary
Caroline Robitaille	Enquête longitudinale auprès des familles séparées et recomposées du Québec / Partenariat de recherche séparation parentale, recomposition familiale	Social Work	Laval
Thecla Schiphorst	Moving stories: digital tools for movement, meaning and interaction	Fine Arts	Simon Fraser
Valerie Steeves	The equality project	Communications / Media Studies	
Rashid Sumaila	OceanCanada	Economics	UBC

Insight Grant Holders

Name	Project	Discipline	Institution
Lisa Dillon	Family reproduction, siblings and the life course: Sibling influences on demographic behaviour in Quebec	Sociology and demography	Montreal
Michelle Levy	Women's print history project, 1750-1830	Literature	Simon Fraser
Jeff Mielke	Data Mining Sound Patterns	Linguistics	NC State University
Allison Muri	The grub street project: topographies of literature and communications in 18th century London	Literature	Saskatchewan
Antonio Páez	Parks, shops and friends: How do the buildt and social environments influence travel behavior	Geography, urban planning	McMaster
Jason Roy	The power of polls - a cross-national experimental analysis of the effects of campaign polls	Political Science	Wilfrid Laurier
Charles Tepperman	Mapping an alternative film history: A database of significant amateur films	Fine Arts	Calgary

Appendix II: Tasks and Deliverables

Tasks

- Provide Portage-related content and participate in a virtual kick-off meeting for the focus group activity (SSHRC will design the program for and host the kick-off meeting);
- Assist the focus group in completing data management plans (DMPs) via the Portage Network's *DMP Assistant*, and provide a verbal briefing on the nature of the assistance provided;
- Develop and administer a questionnaire to learn about focus group participants' experiences in completing the DMPs, including challenges they faced and would expect to face if required to implement the DMPs;
- Design and conduct a day-long capstone event to discuss challenges and support for managing data, and to provide participants with feedback on their data management plans;
- Draft a report on workshop findings, for submission to SSHRC.

Deliverables

- Work-plan – a detailed work-plan for the project, including timeline, staff allocation, and budget within the first two weeks of the contract
- Portage content (e.g. PowerPoint slides) for virtual kickoff meeting
- Verbal briefing on support provided to participants in completing DMPs
- Final questionnaire
- Draft workshop design
- Final workshop design
- Conduct workshop
- Draft report on workshop findings
- Final report

Appendix III: Agenda for the Capstone Event

DMP Workshop Agenda
SSHRC Offices, Room 18-125
350 Albert Street, Ottawa
January 24, 2017

8:30 - 9:00 Registration and Coffee

9:00 - 10:00 Introductions and objectives for the day

- Welcome by Matthew Lucas, Executive Director, Corporate Strategy and Performance, SSHRC
- Workshop participant introductions

10:00 - 10:30 Break

10:30 - 12:00 The Context for DMPs: A review of RDM

- Quick review of research data management (RDM) trends nationally and internationally (C. Humphrey)
- RDM Survey Consortium and Queen's survey results for humanities and social sciences (J. Moon)
- Discussion at tables:
 - How are the trends in RDM changing the way you do research? If they aren't, do anticipate that they will? Why or why not?
 - Do the survey results of the humanities and social science researchers at Queen's resonate with your experiences? If they do, in what way? If they don't, how not?
 - What do you see as the most prominent RDM issue in your research?
- Report back to the full group and discuss the changes needed to move forward from the current situation

12:00 - 1:00 Lunch

1:00 - 2:30 DMPs from the researchers' experiences in this Workshop

- A summary of the DMPs completed for this Workshop (C. Humphrey & J. Moon)
- Survey results from this Workshop (C. Humphrey & J. Moon)
- Discuss at tables:
 - From your experience, what aspects of DMPs do you find useful?

- In what ways are DMPs best utilized by researchers, funders, and other stakeholders in research (please specify who these others are)?
- Report back to the full group: where do we go from here?

2:30 - 3:00 Break

3:00 - 4:30 Integrating DMPs with emerging Tri-Agency data policies

- Review the responsibilities section for researchers in the Tri-Agency Statement of Principles on Digital Data Management (J. Geelen)
- Discuss at tables:
 - What should be done in preparing the next generation of researchers to undertake research data management? In terms of responsibilities, who should be involved in preparing them? What role might DMPs play?
 - What responsibilities do scholarly disciplines have in research data management? How should disciplines address their research data management responsibilities? What role might DMPs play?
- Report back to the full group: where do we go from here?

4:30 - 4:35 Concluding remarks

Capstone Participants / Participant(e)s

Éric Bastien	Social Sciences and Humanities Research Council	<i>Conseil de recherches en sciences humaines</i>
Stephen Batiuk	University of Toronto	
Suzanne Board	Social Sciences and Humanities Research Council	<i>Conseil de recherches en sciences humaines</i>
Ratana Chuenpagdee	Memorial University	
Shannon Cobb	Natural Sciences and Engineering Research Council	<i>Conseil de recherches en sciences naturelles et en génie du Canada</i>
Jacques Critchley	Social Sciences and Humanities Research Council	<i>Conseil de recherches en sciences humaines</i>
Lisa Dillon	Université de Montréal	
Michèle Dupuis	Social Sciences and Humanities Research Council	<i>Conseil de recherches en sciences humaines</i>
Julio Mercader Florin	University of Calgary	
Jeremy Geelen	Social Sciences and Humanities Research Council	<i>Conseil de recherches en sciences humaines</i>
Kaitlin Goertz	Brock University	
Anne Marie Goodfellow	University of British Columbia	
Martin Goyette	École nationale d'administration publique	
Susan Haigh	Canadian Association of Research Libraries	<i>Association de bibliothèques de recherche du Canada</i>
Sara Hormozinejad	University of Calgary	
Chuck Humphrey	Portage Network	<i>Réseau Portage</i>
Vesna Kerezi	Memorial University	
Shahira Khair	Portage Network	<i>Réseau Portage</i>
Mark Leggott	Research Data Canada	<i>Données de recherche Canada</i>
Michelle Levy	Simon Fraser University	
Gary Libben	Brock University	
Matthew Lucas	Social Sciences and Humanities Research Council	<i>Conseil de recherches en sciences humaines</i>
Jessica Mankowski	Canadian Institutes of Health Research	<i>Instituts de recherche en santé du Canada</i>
Lori MacMullen	Canadian University Council of Chief Information Officers	<i>Le Conseil des dirigeants principaux de l'information des universités canadiennes</i>
Karina McInnis	Canadian Association of Research Administrators	<i>L'association canadienne des administratrices et des administrateurs de recherche</i>
Jeff Mielke	NC State University	

Jeff Moon	Portage Network	<i>Réseau Portage</i>
Allison Muri	University of Saskatchewan	
Nicolas Parker	Social Sciences and Humanities Research Council	<i>Conseil de recherches en sciences humaines</i>
Phillippe Pasquier	Simon Fraser University	
Barbara Porrett	International Development Research Centre	<i>Centre de recherche pour le développement international</i>
Caroline Robitaille	Université Laval	
Robbin Tourangeau	Leadership Council for Digital Infrastructure	<i>Le Conseil du leadership sur l'infrastructure numérique</i>
Thecla Schiphorst	Simon Fraser University	
Rashid Sumalia	University of British Columbia	
Charles Tepperman	University of Calgary	
Lori Walker	Canadian Association of Research Ethics Boards	<i>L'Association canadienne des comités d'éthique de la recherche</i>

Appendix IV: A Summary of the Analysis of Submitted DMPs

Data Collection

- The 13 DMPs revealed a complex collection of **raw** and **processed** data types based on:
 - Observational and experimental methods,
 - Quantitative and qualitative measurements,
 - Materials digitized and born digital, and
 - Cross-sectional, longitudinal, spatial, and temporal research designs.
- Some researchers acquired data from an existing source, while others collected new data.
- Several projects produced derived data files through the integration or linkage of other files.
- A variety of file formats are identified across the projects.
 - Thirty-six different file extensions or types were mentioned.
 - On average, six different file formats were used per project.
 - Two projects mentioned 15 different file formats.
 - Common formats were txt, xml, csv, jpg, xls, pdf
- File naming conventions often incorporated content, version, and organizational functions in the names. Databases assigned names to files in a couple of projects.

Documentation and Metadata

- The following types of documentation practices were mentioned.
 - Programming scripts
 - Protocols
 - Self-defined descriptors
 - Database field headings
 - Technical reports
 - User guides
 - Mark-up encodings
 - Transcripts
 - Tutorials, text and video
- The parties most commonly identified as having the responsibility for producing documentation were:
 - a programmer or data manager; or
 - the research team maintaining a shared document.
- Instruments (e.g., a questionnaire) and other supporting materials were often mentioned as part of the documentation.
- Seven of the projects did not use a metadata standard.
- A couple projects used self-defined metadata descriptors.
- Dublin Core and PBCore are standards that were mentioned.

Storage and Backup

- The estimated storage was mostly in the gigabyte range:
 - Less than 50Gb: 7 projects
 - Greater than 1TB: 4 projects
 - TBD: 2 projects
- Many of the projects rely on the backup practices used with university or Faculty IT servers. External drives and researcher workstations were mentioned for backup in two of projects. Library data repositories were also mentioned a couple of times.
- Access to storage was through password-protected methods, including VPN, SSH, or websites.

Preservation

- Most projects are looking at one or more possible solutions. One project will destroy its data five years after the project, according to its ethics agreement.
- Eight projects mentioned consulting with their library about depositing their data with a data repository.
- Two projects are considering using domain repositories.
- Text formats, csv in particular, were mentioned as how data will be prepared for deposit.

Sharing and Reuse

- Several of the projects mention the use of a website to share data. A couple will use a data repository.
- Sharing will be conducted by most projects through a Creative Commons license agreement.
- Two projects mention the Open Data Commons Attribution license (ODC-BY), which is similar to CC-BY but for databases.
- Discovery of a project's data was treated by most as a knowledge mobilization exercise. Only a couple DMPs mention DOI's and library digital discovery methods.

Responsibilities and Resources

- The person or persons responsible for the data during the project were identified to be the PI, a committee, or a programmer or project manager.
 - Seven mention a programmer or project manager.
 - Five mention the PI.
 - One mentions a committee.
- Succession plans for data stewardship were dependent on the number of researchers involved in a project. Teams have options for passing along the responsibility. In smaller projects, replacing the person who leaves serves as the plan.

- Many of the projects will draw upon in kind institutional resources to implement their DMP.
 - Four projects mention Faculty level support.
 - Funded positions were mentioned by four projects.
 - The library was mentioned by three projects.
 - Two mention fees associated with domain repository support.

Ethics and Legal Compliance

- To deal with ethical, legal or IP data issues, many of the projects mention a governance structure, either as part of their project or through their institution.
- Four projects did not see this topic as applicable.
- One project is working on these details.

Appendix V: SSHRC DMP Workshop Survey

Using the DMP Assistant and Interface

The following questions relate to the usability of the DMP Assistant.

Please rate your level of agreement with the following statements.

	Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree	Not Applicable
The registration and sign-in process is straightforward	1	1	<input type="radio"/>	12	<input type="radio"/>
It is easy to make changes to my answers	1	1	1	11	<input type="radio"/>
It is easy to export the completed plan	1	<input type="radio"/>	<input type="radio"/>	13	<input type="radio"/>
The export options include the format(s) I would need	1	<input type="radio"/>	2	11	<input type="radio"/>
It is helpful to be able to add collaborators	<input type="radio"/>	<input type="radio"/>	1	9	4
Being able to share notes with collaborators is useful	<input type="radio"/>	<input type="radio"/>	3	5	6
The DMP Assistant is easy to use	1	<input type="radio"/>	6	7	<input type="radio"/>

Using the Template

Sections found in the Portage template include: data collection, documentation and metadata; storage and backup; preservation; sharing and reuse; responsibilities and resources; and ethics and legal compliance.

2. Was the user interface of the DMP Assistant intuitive?

13 Yes 1 No

Was the user interface of the DMP Assistant intuitive? Please note any issues you encountered.

The length and coverage of the template is what I would expect when documenting a data management plan

9 Yes No 5 Unsure

If no, please comment:

Do you feel that using DMP Assistant prompted you to consider aspects of data management that you would not have included by yourself?

12 Yes 2 No

Please comment:

The questions within the sections are meaningful to my research.

12 Yes 1 No 1 Unsure

If no, please comment:

The questions are comprehensive and don't leave out any data management factors important to my research.

8 Yes 1 No 5 Unsure

If no, what is missing:

The terms, concepts and language used reflect those used in my research area or discipline.

10 Yes 2 No 2 Unsure

If no, what is your suggested terminology:

Were there any questions that you could not answer?

8 Yes 5 No 1 Unsure

If no, please explain why not (e.g. too generic, too specific):

Using the Guidance

Guidance is provided on the right side of the template to help you respond to the data management questions. It can take the form of examples or more detailed information.

The guidance text helped me to understand and answer the questions.

14 Yes No

If no, how can the guidance be improved:

Would more specific guidance be helpful for any of the sections or questions?

	Yes	No	Unsure
Discipline specific	7	3	4
Institution specific	5	5	4
Funder specific	6	4	4

Please comment:

Assistance

In completing your DMP, did you work with the local librarian identified for you?

10 Yes 4 No

If yes, please describe how that relationship was helpful or not in completing your DMP:

If no, please explain why:

Data Management Plans in Policy and Use

The following questions address data management planning from a policy and usage perspective.

Please rate your level of agreement with the following statements.

	Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree	Not Applicable
I already have formal documentation that covers topics addressed in a data management plan	2	7	4	1	<input type="radio"/>
A data management plan is useful to assist in planning a research project (examples of this might include estimating costs, identifying needed supports, allocating resources or	<input type="radio"/>	<input type="radio"/>	3	11	<input type="radio"/>

establishing timelines for specific data actions such as a backup schedule)

Having a data management plan would help me prepare a submission for ethics approval

1 6 4 3

If asked to create a data management plan, a tool with templates and guidance (such as the DMP Assistant) would be helpful

 3 11

Appendix VI: SSHRC DMP Workshop Librarian Questionnaire

	Yes	No	Total
1. Did you meet with this researcher or a designate in person about her or his data management plan?	9	3	12
2. Did you have to initiate the first meeting or did the researcher? [Yes = librarian initiated meeting]	5	4	9
3. Did the researcher have you meet with her or his data manager also?	4	5	9
	1 Time	2 Times	Total
4. How many times in total did you meet with the researcher and/or data manager?	8	1	9

	Not at all well				Very Well	Total
5. Overall, how well do you feel these discussions went?	0	0	1	5	3	9
	Not at all satisfied				Very satisfied	Total
6. How satisfied were you overall with this experience?	0	0	4	5	3	12

	Yes	No	Total
7. Would you do anything differently the next time you are asked to advise on a DMP in the future?	4	8	12

Overall, how well do you feel these discussions went? Please explain.

- I met with the research manager and grad student, both of whom were overwhelmed with the task at hand and not clear why they were being asked to do so much additional work. I spent time explaining the reasoning behind the various steps in the RDM process and how this process might benefit their project. I think they left with a better understanding of the purpose, although not sure that they would be able to complete the tasks with all due diligence in the time frame provided.
- The researcher was open to discussion. We reviewed all the questions together. She mentioned that the whole questionnaire was an opportunity for her and her team to think about all the aspects of data research management before the project take place. We also learned about her experience as a researcher.
- The researcher was really open to discussion and suggestion especially regarding technical details (storage and condition of access during the project), metadata and documentation.
- Note, two participating projects on our campus, only one of the researchers had a meeting to discuss the plan.
- We had a substantial discussion about her data management needs.
- I had a meeting with the researcher and two colleagues of his re: a DMP for their project. To this meeting, I invited a librarian who is a liaison for the researcher's subject area. The liaison and I gave an overview of a DMP and demonstrated DMP Assistant and the Portage DMP template. We were able to address many of the researcher's questions during the meeting. For those questions we didn't have an immediate answer, we followed up with the researcher afterwards. At the time of the meeting, the researcher was in the process of hiring a data manager, and we plan to follow up with the researcher in the new year to meet with the data manager and help set up a Dataverse for the researcher's project. The researcher told us that the meeting was very helpful.
- On my own initiative, I went to see the researcher (who I already knew) and talked with her after a conference she was giving at her department. We talked about the project for about 15 minutes and I told her that I was available to give her a hand when she was ready, if she needed it, but she never contacted me again.
- I met in person with 2 research assistants; the project PI attended via skype. There was some confusion about the goal of the project and the sort of content they were expected to produce. I felt I was able to explain the project and reassure them that there wasn't a right answer for each part of their DMP. It definitely helped to refer to the DMP as a living document they would be able to - and indeed have to - come back and revise because they couldn't have all the specifics from the start.
- I met with the data manager / research manager on this project, who has a good handle on data management strategies within the project and confident about how to write the document. He requested more information about what preservation services would be available through the library and other hosting possibilities, I was able to direct his inquiries, but unable to offer definitive answers in regards to these questions.

What were the primary reasons that you had no contact with the researcher about her or his DMP for this Workshop project?

- I'm unsure--I reached out to the researcher at the beginning of the pilot project offering support in any way, shape or form, but heard nothing from them.

- I did not actively pursue the researcher - I reached out via email towards the end of the project period as reminder that I was available. I would like to assume that no response meant that no additional support beyond the tool was required.
- I did not actively pursue the researcher at the start of the project and the researcher did not contact me for support. I did check-in via email towards the end of the project timeline to again put my contact information in front of the researcher.
- I offered help but the researchers did not have time to meet with me.

Is there anything that you would like Portage to help you with in the future in advising on DMPs?

- I think we have a lot of work to do as a community to develop informational resources and exemplars to serve as a knowledge base--I think we'll get there in time, and that this pilot is certainly a step in that direction.
- I think the DMPTool is very approachable.
- Not really. The biggest challenge at my institution is the lack of cohesive institutional infrastructure and support. We can tell researchers how they *should* be doing things ideally, but cannot provide them with concrete solutions to accomplish them.
- With all the resources that PORTAGE already provide, we feel that the ball is now in our hands... By the way, this project to bind us with a researcher in a real context was after all a great idea by PORTAGE.
- Storage and preservation of data outside our institution.
- Nothing comes to mind.
- No
- It would be great if Portage could inform us about any upcoming changes to DMP Assistant (e.g., interface, new templates).
- It might be useful to have sample DMPs from Canadian researchers in different disciplines.
- Don't know yet
- I'm completely unprepared to answer questions about the cost of data management. It's so specific to projects and institutions, I'm not sure how Portage can help, but if you can, that's what I could use!
- If SSHRC (or any one funding agency) has particular requirements within the DMP, either in terms of format or information.
- An option in the DMP Assistant for library employees to act as a reviewer with the press of a button would be great!

Would you do anything differently the next time you are asked to advise on a DMP in the future?

- Perhaps I would follow-up with them once more, but I'm a bit ambivalent about it. I don't want to send unwanted emails to the researcher and I would like to establish a relationship where they have a responsibility to contact me when they need assistance--rather than waiting for me to offer the invitation once more.
- More time invested in connecting with the researcher about their research versus only offering to support the planning aspects.
- A group has just formed on campus with representation from various units (ITS, graduate studies, office of research, CIO, etc.). I hope in the future to be able to tap into the expertise of the

members of this group. It will be some time, though, before our campus is ready to truly support RDM.

- Being more prepared when it comes to giving advice for preservation and sharing of data (what platform is more appropriate given the type of data produced during the project?).
- Most of the difficult to answer questions were very situation/discipline specific.
- A "not sure" bullet would've been more appropriate above. I looked up the researcher's research on her personal site and elsewhere before I met with her. I'm not certain this would be possible in every case but I will definitely email for at least some preliminary information before meeting so as to be better prepared. This is very analogous to a 'traditional' reference interview.
- I think that the referral system worked well. I appreciated that SSHRC/Portage encouraged researchers to seek DMP help from libraries and librarians. Researchers may not think of libraries as the first place to seek such help.
- I would be more proactive in seeking to set up a meeting to review the researcher's project - while I assume the researcher did not contact me because they felt prepared and able to manage the DMP, I cannot be certain.
- After contacting a researcher twice to offer my help, I did not feel comfortable insisting any further
- Discuss early on a timeframe or schedule for checking back regarding the items in the plan that can't be answered right away. There are a few important tasks that are described in the plan as being "in progress" or "under development" and involve discussions with library personnel; it's easy to imagine those being forgotten if the plan isn't revisited regularly.
- I feel I answered the questions and directed further questions appropriately. The patron left feeling comfortable with how to write his document.
- I offered help but the researchers completed the DMP on their own. This speaks to how easy and helpful the tool is to use.

Appendix VII: Capstone Evaluation by Researchers and Observers

Capstone Evaluation by Researchers and Observers						
1. Did the capstone event meet your expectations?						
Scale	Failed expectations	Met some	Met all expectations	Exceeded expectations	No opinion	Total
Researchers	0	6	6	0	0	12
Observers	0	0	4	1	0	5
2. How important to you was the information at the event about DMPs and data policies?						
Scale	Not at all important	Not very important	Important	Very important	No opinion	Total
Researchers	0	1	9	2	0	12
Observers	0	0	1	3	1	5
3. How satisfied were you in having an opportunity to express your thoughts during the event?						
Scale	Not at all satisfied	Somewhat satisfied	Satisfied	Very satisfied	No opinion	Total
Researchers	0	1	7	4	0	12
Observers	0	0	0	4	1	5
4. Overall, how would you rate the event?						
Scale	Very poor	Poor	Good	Very good	Undecided	Total
Researchers	0	1	7	4	0	12
Observers	0	0	1	4	0	5

Seven respondents provided comments at the end of the questionnaire.

1. Excellent activity and valuable input from researchers and stakeholders. Through this capstone discussion, SSHRC has effectively demonstrated that the views of the

researchers and the broader stakeholder community are important to guide the development of data management policy and practice for Tri-Council-funded research. SSHRC is to be commended for leadership in this area. Bravo!

2. Je crois que les sessions de travail en sous-groupe auraient été davantage efficaces si une personne du CRSH avait eu le mandat, dans chaque sous-groupe, de veiller à ce que les participants répondent aux questions posées et ramène au besoin les participants.
3. It seems crucial, in my opinion, that SSHRC capitalizes on the momentum and goes from discussion to policy in order to ensure that all tri-council applicants detail what the nature of their data management plan will be and that educational activities increase at the conference, workshop levels. Most of our colleagues have a perfunctory understanding of data management, if any at all. Count on me, if needed.
4. Flying people for so little is a waste of public money. Sshrc and other funding agency need to learn how to work online, make video presentations and share information and gather feedback in efficient ways. It also seems like a bad idea to let University and especially libraries to handle this in a decentralized way...For once, a national system would make sense.
5. The continuation of this training program can be very beneficial to our research project.
6. I enjoyed the presentations and the ability to work in small groups on focused questions. There was a good representation of different stakeholder groups; and an openness to different points of view. Overall, very well organized and interesting. thank you.
7. Great initiative overall.

Appendix VIII: Notes of Audience Discussions at the Capstone Event

8:30 - 9:00 Registration and Coffee

9:00 - 10:00 Introductions and objectives for the day

Researchers at the event were asked to identify themselves and briefly describe their impressions of using the DMPs in their work.

- Overall, researchers expressed appreciation for the DMP process and the opportunity to reflect upon their various research and data collections processes.
- Common themes among comments included that some of the work involved was already being done in some capacity, and that there was appreciation for the DMP process for providing the opportunity to incorporate these various actions into a concrete process. One challenge encountered by several researchers involved dealing with the idiosyncrasies of researcher-specific and subject-specific data. Another common challenge identified was the amount of work involved in creating the DMP, which tended to expand beyond creating the DMP itself.

10:00 - 10:30 Break

10:30 - 12:00 The Context for DMPs: A review of RDM

Quick review of research data management (RDM) trends nationally and internationally (C. Humphrey)

- General comments received at the end of the presentation centered on constraints and challenges in the vertical organization of funding structures and data management requirements. Participants were unsure how to work across interdisciplinary boundaries and how to foster collaboration between researchers in different fields.
- The CIHR training modules currently in development were mentioned. Participants asked whether these modules would be open to SSHRC or other funding agencies and were told that they would be made public on the web by CIHR, hopefully by the fall.
- Another question posed was whether SSHRC grant reviewers would be looking for elements of DMPs in applications. Participants were informed that there was no formal requirement for DMPs yet. For the moment, SSHRC is mostly focused on capacity building and wants supporting tools widely available to applicants before requirements are considered.
- A comment was raised that due to the increase in interdisciplinary projects, there exists a real need for structures to enable the Tri-Councils to work together and use the same language. In response, participants were told that SSHRC was aware of the need to collaborate more effectively across the Tri-Agencies and that the common Statement of Principles is an example of this awareness.
- A final comment raised the impact that vertical structures for funding models have on the horizontal management of data across researchers and institutions, as well as the intergenerational management of data, since there's no way of knowing who in the future will be using the data that researchers are currently creating.

RDM Survey Consortium and Queen's survey results for humanities and social sciences (J. Moon)

- One of the Queen's survey questions showed confusion about what data repositories are. The capstone participants expressed concern over their own general lack of awareness of subject-specific repositories.
- When queried, most researchers in the room agreed that they brought their own research into their pedagogical practices.
- One participant asked when the RDM Survey Consortium's data will be made available to the public. J. Moon responded that currently there is a link on the Queen's website to the Engineering

and Science report and that the Social Sciences and Humanities report is forthcoming. The plan is for other participating universities to release their work in a similar way. There is also a plan to produce an aggregated data file across all participating institutions, although there may be a need to suppress certain variables because of privacy considerations. Portage will be serving as the clearinghouse for this Consortium's work. For institutions interested in participating in the survey, the Research Intelligence Expert Group will be able to assist them.

Discussion at tables:

- *How are the trends in RDM changing the way you do research? If they aren't, do you anticipate that they will? Why or why not?*
- *Do the survey results of the humanities and social science researchers at Queen's resonate with your experiences? If they do, in what way? If they don't, how not?*
- *What do you see as the most prominent RDM issue in your research?*

Summary of comments at the first table to report back:

- The discussion focused on the change that researchers aren't just producing outputs for themselves but are also expected to share their data for the use of others.
- Important questions that were raised included when and how to share data and at what point in the research lifecycle should data be preserved. While the focus may primarily be at the end of the research process, participants wondered whether mechanisms were available to make data accessible to others during the data gathering process, which may occur over an extended period of time.
- In regards to preservation, participants felt that the form of their data was an important component that has not been sufficiently addressed. The question of how to preserve presentation forms, environments, and experience was an important concern, in addition to preserving the actual data.
- Another concern expressed was whether potential users had sufficient skills to assemble the flattened data from a relational database into a usable product.

Summary of comments at the second table to report back:

In response to Question 1:

- Participants were encouraged to see how trends in research are changing. For example, the push for open access and an increased emphasis towards sharing is changing the ways in which researchers conceive of research processes and outputs.
- Participants noted that the division of labour in research projects is being impacted. In many cases projects require data managers, or at least access to individuals with skills for managing and sharing data. Noticeable impacts are being felt by their graduate students, who need more training with RDM.
- Another trend observed involves the relationship between planning and conducting research. Planning is much more involved at the front end now, so considerations need to be made well before data is collected on the ground.

No discussion was held in response to Question 2.

In response to Question 3:

- Participants identified access to sufficient resources as being the most prominent issue, particularly around capital and human resources. Institutions were considered important factors in providing access; for example, access to institutional repositories and knowledgeable data librarians. However, even these may be insufficient. For example, institutional repositories may not be appropriate for certain types of data and librarians typically provide support at the front end of research projects when support may be needed throughout the entirety of a research project.

- Another challenge that was discussed involved issues in partnering with industry to conduct research and how to publically share those outputs.
- A large issue is that researchers are not being recognized fairly by the academic rewards system for producing and sharing their data.

Summary of comments at the third table to report back:

Discussion at this table was restricted to Question 1.

- Participants at the table were divided over whether they felt that a DMP changed their approach to research.
- Participants noted that trends in RDM posed challenges for interdisciplinary research groups, impacting how they share data in various formats, which may not be compatible, and how they agree upon issues of data ownership and stewardship.
- One participant raised concerns regarding the misinterpretation of publically accessible research data and questioned whether the potential misuse of their data was valid.
- One researcher working with indigenous subjects expressed concern over differing notions of ownership, sharing, and access, which they felt needed to be addressed.
- Researchers want to receive credit or rewards resulting from RDM and anticipate that a shift in the academic rewards system would encourage change to common practices. They felt that enthusiastic adoption of RDM will likely only occur if the academic rewards system drives the process forward.
- Participants expressed desire for greater cohesion between requirements of their local Research Ethics Boards and the Tri-Councils. They felt that some rules were very confusing and in some instances, they bump into each other; for example, different ideas concerning secondary use of data and differences between anonymous vs anonymized data.

12:00 - 1:00 Lunch

1:00 - 2:30 DMPs from the researchers' experiences in this Workshop

A summary of the DMPs completed for this Workshop (C. Humphrey & J. Moon)

Survey results from this Workshop (C. Humphrey & J. Moon)

- Comments received at the end of these presentations surrounded topics of preservation and the capabilities of repositories.
- One participant asked how researchers should deal with the fact that datasets are living entities. Once in a repository, how should data sets be kept up to date and how should access evolve with a project? The presenters spoke about the use of versioning to identify datasets that change over time. At some point in time, a snapshot of the data is taken and processed and this practice then continues into the future of a project.
- Participants asked if there is any benefit to using more than one repository for their data. Presenters explained that in a federated system, data could be discovered across multiple repositories and that Portage and Compute Canada are in the midst of developing a Federated Research Data Repository, which would help with this question. In the absence of this platform, researchers could use a staging repository to hold their data until such infrastructure becomes available.
- Participants commented on the challenge of remaining linked to data once it is in the public domain. Presenters noted that issues underlying data provenance continue to exist. However, there is the potential for block chain technology to forever identify researchers with their data.

Discussion at tables:

- *From your experience, what aspects of DMPs do you find useful?*
- *In what ways are DMPs best utilized by researchers, funders, and other stakeholders in research?*

Summary of comments at the first table to report back:

- Participants discussed the concept of active DMPs being used throughout the research lifecycle, rather than something that's done only once at the beginning of a project.
- One participant mentioned having created tools in order to access data from their DMP for reuse, and wondered whether similar tools were available. C. Humphrey responded that Portage is working on APIs to exchange information from DMPs to relevant stakeholders.
- The group identified other useful features from DMPs, including the development of an underlying metadata model that identifies data within them and the assignment of DOIs to published DMPs, which would allow them to be shared and educate other researchers.

Summary of comments at the second table to report back:

- Group participants discussed the division between archival data and live data. Comments expressed that DMPs seem to focus primarily on the former, whereas researchers focus primarily on the latter, which evolves and must be updated.
- Participants asked where the resources and support are to manage live data. Discussion then centered on the importance of versioning datasets and taking snapshots of data sets at various milestones or points in the research lifecycle.
- Uncertainty of when in the research lifecycle to share DMPs was also expressed. For example, is it only shared once a grant is accepted?
- The concept of “designated communities” was raised, which describes the potential future user base who a researcher might assume to be working with their data in the future. When preparing a DMP, it is being done also with this intended user base in mind.

Summary of comments at the third table to report back:

- Participants reported that one useful aspect of the DMP process was that it triggered their thinking about the transferability of their data and succession planning.
- Conversations were also raised around the various people responsible for RDM and for completing DMPs. Participants expressed that it would be desirable for more guidance about who should take responsibility for various DM roles and DMP components.
- Participants discussed where they found support in completing their DMPs. The general consensus was that it was a mix between their institutional libraries and IT departments. Although participants all felt that this approach may not be possible for all institutions.
- The concern about how students should be involved in the DMP process was raised, particularly since they are not often connected to a project for long durations.
- Can a researcher not consider some data in the DM process, i.e. does all data associated to a research project need to be managed, preserved and made accessible?
- Rights surrounding data were discussed. Participants expressed uncertainty about what data rights to assign to their data and requested more guidance and standard knowledge when selecting from the options being presented.
- Participants also acknowledged that there may be some bias associated with the researchers in the room regarding their opinions on the openness of data.
- Finally, participants saw a DMP as helping justify funding for RDM staff. Anecdotally, completing the DMP helped one researcher form an argument for services that their institutional IT staff should provide.

2:30 - 3:00 Break

3:00 - 4:30 Integrating DMPs with emerging Tri-Agency data policies

Review the responsibilities section for researchers in the Tri-Agency Statement of Principles on Digital Data Management (J. Geelen)

Discuss at tables:

- *What should be done in preparing the next generation of researchers to undertake research data management? In terms of responsibilities, who should be involved in preparing them? What role might DMPs play?*
- *What responsibilities do scholarly disciplines have in research data management? How should disciplines address their research data management responsibilities? What role might DMPs play?*

Summary of comments at the first table to report back:

- One topic of discussion concerned the adjudication of DMPs and how it would impact judgement of research projects, particularly regarding funding decisions. If funding decisions are contingent on DMPs, participants felt that this could be problematic since they were unsure who exactly would be qualified to pass judgement. A related concern was whether a requirement for DMPs might skew funding towards more data-driven subjects and away from less data-intensive research. A suggestion was made to treat DMPs similar to the requirement for ethics approval, which is only required for some projects and doesn't impact funding decisions.
- Another topic mentioned, but not expanded upon, raised the question about which level of scholarly organizations are best equipped to assist with RDM processes.
- Regarding institutional requirements for RDM, to avoid confusion initially, participants proposed that institutions begin with simple and straightforward requirements, then over time progress in complexity.
- To prepare future researchers, participants felt that questions about what constitutes research data and who makes that decision should be resolved. By extension, what areas of research should require DMPs. Participants felt that these questions were particularly problematic for the social sciences and humanities, where the language of data has not been established. Much work lies ahead in articulating this to researchers.
- A question was raised whether the analysis of existing information constituted the creation of data, for example, the use of archival resources. One participant expressed concern that there may be a skewing of language in the humanities towards framing research topics in "data terms".
- The availability of an opt-out option from the DMP was considered for researchers who don't consider their work to involve creating data. This point was opposed by several participants in the room who argued that there will likely be benefits to all researchers in completing a DMP, even though the process is likely to be alien to many humanists.
- The point was made that a benefit of the DMP and of making the outputs of research open and accessible are the opportunities these actions create for others, which occur regardless of how data-intensive the area of research may be.
- R. Sumaila described Oceans Canada's very loose definition of data and noted that his research group's Atlas includes First Nations narratives as data.
- It was described that at SSHRC, similar discussions have been held with researchers who are concerned about the changing culture in their community. However, the main challenge lies in showing researchers what aspects of RDM apply to their work. The relationship between policy change and uptake is complex and there is usually a lag of several years before motivation is evident in the affected community.

Summary of comments at the second table to report back:

- Participants stressed that work at the institutional level must be done to prepare the next generation of researchers and that the Tri-Agencies and research communities have work ahead of them in articulating RDM.
- With respect to DM planning, participants noted that funding remains an important issue. Different research groups have various funding levels and thus various capacities to hire data managers. Despite this variation, the pressure is being felt across all groups.
- The subject of the peer-review process of DMPs was discussed. Particularly, how a review process might take shape considering the diversity of approaches to RDM adoption in

communities. Complicating this process, RDM is not a linear process and it may therefore be challenging for some researchers to incorporate a full DMP into their grant applications.

- Discussions were held about how to bring about a culture change and advance RDM with the next wave of researchers. Participants emphasized motivating researchers about the value of their data as a research output.
- The role for champions to promote RDM in their communities was discussed. For example, participants spoke highly of CIHR's effort to provide training for their research communities.
- Participants also discussed the possibility of them inviting change. As they too are apart of the culture of their research communities, their own actions can promote RDM. Education and learning opportunities were seen as valuable methods for illustrating the meaning of RDM, fostering curiosity, and opening space for culture change.

Summary of comments at the third table to report back:

- Participants expressed a need to provide education to graduate students, but also train researchers. They noted however the lack of tools and training opportunities available to them.
- Related questions arose around how to incentivise training. For example, would it become a requirement for researchers at an institutional or organizational level. Could a system of credits or certification be developed to recognize researchers who had taken the time to train themselves in RDM.
- Participants emphasized that a smooth transition to requiring DMPs is needed, and that further guidance for institutions and researchers on RDM is required. A phased approach was proposed, with only funded projects requiring DMPs at first before broadening requirements.
- The question was posed whether institutions that could not provide support for RDM -- in terms of education, assistance, or storage -- would be penalized. Participants agreed that even these smaller, less equipped institutions should be given the support to comply with future requirements for RDM. In the eventual adjudication of DMPs, it is hoped that the granting bodies would take into consideration whether applications were being sent from smaller institutions.
- Regarding the adjudication of DMPs, participants questioned whether this action might best be left to the Research Ethics Boards.
- In response to these comments, C. Humphrey informed participants that conversations are taking place with the organizations that support research on campuses. One of Portage's guiding principles is to create a level playing field for all Canadian institutions.

4:30 - 4:35 Concluding remarks

- Participants were informed that there would be opportunities to provide feedback on policy development. They were encouraged to go back to their institutions as ambassadors and let their colleagues know how important this issue is to them, and that the granting agencies are taking this subject seriously.
- A framework for SSHRC's thinking about RDM going forward was presented:
 - This work is not about SSHRC, but about better research going forward.
 - SSHRC is not on the search for perfection and recognizes that the process will need to be gradual. SSHRC also recognizes that there are gaps in the current system, as well as differences between researchers and institutions, so they are not looking for a one size fits all response. Thus, they will be looking to the research community to inform their future developments.
 - SSHRC acknowledges that there are differences between managing and sharing data. However, they recognize that the reason for managing is partially for sharing, although that there are still good reasons for managing if sharing is not possible.
 - A multidisciplinary, cross-agency approach to RDM is required. Thus, SSHRC is working with its sister agencies who also recognize that a joint approach is best and that consistency across existing policies is important.
 - SSHRC is currently grappling with questions over adjudication and any decisions will

likely not be set in stone. They are looking for feedback on the involvement of research ethics boards and local research offices in playing roles in this process, as well as on the rolling out of a phased approach to implementation.

- o Participants were reminded that funding is available to support getting research communities together to discuss RDM topics, for example Connection grants.