

## **DMP title**

**Project Name** My plan (Portage Template)

**Principal Investigator / Researcher** Sean Sears

**Description** It's good if you can summarize the main idea of your research!

**Institution** McMaster University

## **Data Collection**

**What types of data will you collect, create, link to, acquire and/or record?**

Audio recordings of interviews with relevant experts will be created. From those recordings, transcripts will be typed out by a select number of predefined individuals who are active researchers on the project. From these transcripts, numerous qualitative outputs, including qualitative analysis techniques and reports, will be developed/generated.

**What file formats will your data be collected in? Will these formats allow for data re-use, sharing and long-term access to the data?**

The audio files will be recorded as .mp3 files.

The audio files may be deleted after two years, there is zero intention for this files to be shared outside of the listed researchers on the ethics documentation.

The transcripts will be typed out in microsoft word and archival copies will be kept as text files.

The qualitative analysis will be conducted in nvivo, staggered and final outputs can be exported as csv files, but will review will otherwise require a copy of nvivo to have the same level of functional interaction.

**What conventions and procedures will you use to structure, name and version-control your files to help you and others better understand how your data are organized?**

Audio recordings and transcripts are named by the participant / organization represented and the date.

Final nvivo outputs are the project title + QA + Vx.x

Inside of Nvivo, copies of the project can be made as various steps within the coding process. At each the line, focused, and axial coding steps, the sub-folders holding the analysis within nvivo, are named as such. If multiple versions of the coding are generated, a version number would be added.

## **Documentation and Metadata**

**What documentation will be needed for the data to be read and interpreted correctly in the future?**

Transcripts are full scripts of the conversation. The final reports developed would help a third party to understand the context.

The memoing associated with the coding process, if written down, would help to explain how the coding process took place. Again, the final report would be helpful as it discusses the coding process and steps, and identifies the key terms. Otherwise the basis of the analysis is grounded in the interview text.

**How will you make sure that documentation is created or captured consistently throughout your project?**

Audio recordings are recorded from the first question until the conversation is completed.

Transcripts are verbatim and are checked by two researchers for their accuracy.

The coding process is completed by multiple researchers who compare and contrast independent coding results through a comparison process, easily enabled by nvivo. This process helps to ensure that researcher bias is limited and interpretation of the raw results are similar. Documentation between how codes are developed are kept.

**If you are using a metadata standard and/or tools to document and describe your data, please list here.**

Question not answered.

**Storage and Backup**

**What are the anticipated storage requirements for your project, in terms of storage space (in megabytes, gigabytes, terabytes, etc.) and the length of time you will be storing it?**

Storage space will be in the tens of megabytes range, not expecting to exceed 500MB.

Audio will be kept for at least two years.

Transcripts will likely be kept indefinitely.

Nvivo and other coding outputs will likely be kept indefinitely.

**How and where will your data be stored and backed up during your research project?**

Audio recordings are kept on the institutes private server and is accessible only by authorized users.

Transcripts are kept on the aforementioned server. Copies of which will be kept on the primary researcher's encrypted computer for analysis.

The PR's computer is backed up daily to both a portable and desktop extenal hard drive. Cloud copies are kept on the institute's remote-access server.

**How will the research team and other collaborators access, modify, and contribute data throughout the project?**

Original documents are accessible via the server.

Transcripts are not to be modified.

The nvivo coding is completed through the PR's computer. Only one user can actively

edit the data at a time on a local machine. Transferring copies of the nvivo code files will be completed via the server, wherein an up-to-date copy of the file will be uploaded to the server, an authorized user can then check out the file from the server (to note to other authorized users that the data is actively being used/edited).

The research data is not to be transferred via email due to the confidential nature of the data.

## **Preservation**

### **Where will you deposit your data for long-term preservation and access at the end of your research project?**

The institute's private server maintained by University administrators.

### **Indicate how you will ensure your data is preservation ready. Consider preservation-friendly file formats, ensuring file integrity, anonymization and de-identification, inclusion of supporting documentation.**

Long-term storage copies of the data will be made and kept as .txt and .csv to ensure maximum preservation. Audio as .mp3

The audio and transcripts are inherently identifying and are not to be shared outside of the research project, including the initial coding documents. Higher level coding outputs will be available as a csv document.

Given these are text files, quality degradation is not a concern. The nvivo software has an export wizard; it is the role of the researcher to use this tool to select .csv as the output format and the appropriately name the file.

## **Sharing and Reuse**

### **What data will you be sharing and in what form? (e.g. raw, processed, analyzed, final).**

Analyzed and final data will be shared. Only anonymized processed data may be made available.

EDIT:

As per the ethics requirements associated with this research project, the transcripts are considered to be identifiable materials as they are not anonymized, nor is there a reasonable manner in which they could be anonymized, and as such, they will not and cannot be shared outside of the primary researchers identified in the ethics application.

### **Have you considered what type of end-user license to include with your data?**

EDIT:

As per the aforementioned, a EULA is not needed in respect to the transcript data.

### **What steps will be taken to help the research community know that your data exists?**

Given the sensitive nature of the data and the ownership of the data (a semi-private

research institute) the intent is to not share the data except for the final outputs.

EDIT:

By virtue of the paper being written and published, the research community will know that the interviews took place and given the research methodology will know that transcripts were created in order to facilitate the analysis. The higher level coding results will be included as a table in the research output.

## **Responsibilities and Resources**

**Identify who will be responsible for managing this project's data during and after the project and the major data management tasks for which they will be responsible.**

The primary researcher is directly responsible for managing the data during the research process. After the completion of the research project and the final data is deposited to the server, it is the responsibility of the institute to maintain the security of the data. It is the PR's responsibility to maintain the security and redundancy of the data of their workstation, it is the responsibility of the institute to maintain the redundancy and security of the data deposited on the server.

**How will responsibilities for managing data activities be handled if substantive changes happen in the personnel overseeing the project's data, including a change of Principal Investigator?**

Back up copies of all data are kept on the institute server. Given major changes, the most recently deposited data and associated information would be available. The memoing associated with the coding process would guide an individual if they were to over take the process mid way.

**What resources will you require to implement your data management plan? What do you estimate the overall cost for data management to be?**

All equipment required (server, computers, hard drives, network communication) is already owned and the additional cost is only measured in the hours of time required to maintain the plan.

## **Ethics and Legal Compliance**

**If your research project includes sensitive data, how will you ensure that it is securely managed and accessible only to approved members of the project?**

Data is formally kept on the institute's private server, which is only accessible by authorized users, as is the associated web space.

The PR's computer and associated hard drives are password protected and encrypted; can be wiped remotely if the computer is lost or stolen.

**If applicable, what strategies will you undertake to address secondary uses of sensitive data?**

Question not answered.

**How will you manage legal, ethical, and intellectual property issues?**

Data that is not anonymized is not to be shared out side of the research group.

This document was generated by DMP Assistant (<https://assistant.portagenetwork.ca>)