
2021 Redux Survey of Visible Minority Librarians of Canada

A Data Management Plan created using DMP Assistant

Creator: Yanli Li

Principal Investigator: Yanli Li

Data Manager: Yanli Li

Affiliation: Wilfrid Laurier University

Funder: Digital Research Alliance of Canada

Template: Portage Template

Project abstract:

The Visible Minority Librarians of Canada (ViMLoC) Network conducted the first comprehensive survey on visible minority librarians in 2013. ViMLoC conducted this survey again in 2021 to examine the changes in the library landscape with regards to visible minority librarians in various librarian positions including leadership. This survey will help ViMLoC identify the needs of visible minority librarians and propose projects to empower them in their current positions or their future career development such as mentorship programming, leadership training, and networking opportunities. The data can be also useful to library administrators, librarians, and researchers working on multicultural issues, diversity, recruitment and retention, leadership, library management, and other related areas.

This data management plan was created by the 2021 ViMLoC Redux Survey team: Yanli Li, Maha Kumaran, Allan Cho, Valentina Ly, Suzanne Fernando, Michael David Miller, and Michael Steeleworthy, Data Librarian at Wilfrid Laurier University.

Last modified: 12-06-2023

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

2021 Redux Survey of Visible Minority Librarians of Canada

Data Collection

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

Participants in this research will be visible minority librarians in Canada. [The Canadian Employment Equity Act](#) defines visible minorities as "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour". The visible minority population consists mainly of the following groups: Chinese, South Asian, Black, Arab, West Asian, Filipino, Southeast Asian, Latin American, Japanese and Korean, and the intersectionalities of these ethnic groups. There may also be other mixed races such as White, Indigenous, and visible minorities that qualify as visible minorities. Indigenous, White or Caucasian only librarians will be excluded.

We will be collecting survey data using Qualtrics through Wilfrid Laurier University. Participants will respond to multiple choice questions and open-ended questions. Data will be collected on visible minority librarians' demographics, education, and employment situations.

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 survey data from Qualtrics will be exported to MS Excel spreadsheet and CSV format. Aggregated data and charts from Qualtrics can be exported as MS Word or PDF format.

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?

The PI will create three folders to organize data files on her institutional OneDrive which is password protected. A file naming convention will be used to name the various data files within each folder. Each file will be named with a short description/acronym to reflect its content, followed by the date of creation, and version.

Raw data folder: The exported survey data from Qualtrics for each respondent will be named as rawdata_20210401_v01, which represents the raw data exported from Qualtrics on April 1, 2021, version 1.

Cleaned data folder: The raw data will be deidentified (removal of some geo-coded data including IP address, latitude, and longitude, and duplicating responses), and renamed accordingly. For example, cleandata_en_20210503_v01, which represents cleaned data for English survey created on May 3, 2021, version 1. For French survey, "en" is replaced with "fr".

Aggregated data folder: there will be two subfolders.

Subfolder "Aggregated report_Qualtrics" will include aggregated data exported directly from Qualtrics. For example, file name Aggregp_en_20210503_v01.doc represents the report for English survey exported from Qualtrics on May 3, 2021, version 1. This report includes the number and percentage of respondents in response to each question, text responses for the "other" field of some questions, and answers to open-ended questions.

Subfolder "Processed data" will include the data that may be created in the process of analysis, including visual files and graphics. The file name can be specific to reflect its content. For example: academic_en_20210503_v01.docx represents the English survey data on academic librarians only, created on May 3, 2021, version 1. The basic principle is to have file names consistent and make them easier to understand and share within the team.

Documentation and Metadata

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

We will provide a description of the project, the methodology, data collection methods, and information on data de-identification and analysis.

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

We will have written protocols for documenting the survey. The questionnaire and Consent Informed Letter are available.

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

We will have a data dictionary to define different data fields and types. We will provide explanations on data coding.

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?

We expect to have less than 3 gigabytes of data for the survey.

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

The Qualtrics server is in the US and raw survey data will be stored there. All exported survey results will be stored on the PI's password-protected institutional computer.

The PI will provide access to cleaned individual data to two other Co-I's and aggregated data to all collaborators through her institutional OneDrive which is password protected. Identity management is mediated through OneDrive and university protocols for security purposes.

The identifiable data results will be safely destroyed after 5 years post publications following the PI's institutional protocols.

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

The PI has access to the Qualtrics survey results. She will share the results via OneDrive that is password protected. Raw data will not be downloaded directly from Qualtrics by collaborators in this project.

Preservation

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

We can deposit the questionnaire and results summary report to Wilfrid Laurier University's Dataverse.

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.

Data will be converted to .csv format. All graphs if any will be converted to .svg formats for long-term preservations. Any accompanying documents will be converted into .txt format or markdown format.

Sharing and Reuse

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

Analyzed data will be shared with the public through presentation or publication.

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

We will use [Creative Commons BY-SA-NC 4.0](#) for our dataset.

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

We will link our dataset to all our publications and presentations arising out of this study. We will also have a DOI associated with the data so there is permanent and persistent identification to data through the Wilfrid Laurier University Dataverse.

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 PI is responsible for the management and security of the data during the life of the project. Decisions for access, sharing, and use will be made collaboratively with the team. After project completion, the PI and team will be responsible for determining what data is to be archived and what data is to be disposed of. Decisions will be made in accordance with obligation to research subjects and the team's respective REBs.

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?

We do not anticipate any changes. There is more than one researcher in this team who are knowledgeable about handling the data management plan (DMP). If necessary we will seek advice from other data experts who work in equity, diversity and inclusion research areas in librarianship.

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

This project is expected to require resources already available to all team members and no additional costs are expected through the life of the project. The PI will put all datasets and files into her institutional OneDrive and share the necessary data with her team members. The team expects to use the Scholars Portal Dataverse service for long-term deposit of its questionnaire and results summary report, all of which have no additional costs.

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?

This project includes survey data with human subjects, with methods and an approach approved by Research Ethics Board (REB) at Wilfrid Laurier University, University of Saskatchewan, and the University of British Columbia. All sensitive data associated with the project will be secured on Laurier's campus network, which meets the particular needs of this special data. Only the PI will have access to raw data. The two Co-Is who require access to the cleaned data will have their access facilitated through OneDrive, of which Laurier and the two institutions are a part of. If they choose to download the data for easier data editing and processing, the two Co-Is will secure the data on their password-protected computers. Other team members who require access to aggregated data but are not at a Microsoft institution will still be able to get access by obtaining a free Microsoft account.

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

We do not anticipate sharing any identifiable data with other researchers for secondary use.

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

N/A