This Research Data Management Toolkit was developed by the Collaborative Research Center for American Indian Health (CRCAIH) and the CRCAIH Tribal Partners. The toolkit is intended to serve as a resource for American Indian/Alaska Native (AI/AN) Tribes or other Indigenous Nations developing policies around research data management and is available free of charge at www.crcaih.org.

Since 2012, CRCAIH has partnered with tribal communities and health researchers within South Dakota, North Dakota, and Minnesota with a goal to build tribal research infrastructure and transdisciplinary research teams focused on improving American Indian health through examination of social and environmental influences on health.

Correspondence, including questions and requests should be directed to info@crcaih.org

Acknowledgements:

We would like to acknowledge the CRCAIH Tribal Partners and thank them for their input; without them, this toolkit would not have been completed:

- Fond du Lack Band of Lake Superior Chippewa
- Oglala Sioux Tribe
- Rosebud Sioux Tribe
- Sisseton-Wahpeton Oyate
- Tribal Nations Research Group, Turtle Mountain Band of Chippewa Indians

We would like to also acknowledge the CRCAIH staff who contributed to this document: Michaela Seiber, Melissa Buffalo, DenYelle Kenyon, Jessica Heinzmann, Tess Weber, and Jessica Hanson; Yvette Roubideaux (NCAI Policy Research Center).


This work was supported by the National Institute on Minority Health and Health Disparities of the National Institute of Health under award number U54MD008164.
WHY DEVELOP A RESEARCH DATA MANAGEMENT PLAN?

What is Data Management?

Research studies often result in a significant amount of data that is collected from participants. This data could be qualitative data, including transcripts, notes, and database entries, and quantitative data, including spreadsheets, statistical software program data sets, or even bio-specimens. Once the data has been analyzed and results reported, tribes are increasingly requiring researchers to return the data to the tribe in whatever format it is in, whether it is raw data, or data that has been analyzed. Sometimes, the same or different researchers may want to use that data for additional research studies. Once the tribes approve any new projects, they will then need to provide the data again to the researchers. The tribe will need to have a data management plan in place to be able to handle data that is returned to them and, if they choose, to provide the data to researchers for new research projects.

Why Do Tribes Need To Think About Managing Research Data?

Building tribal research capacity and infrastructure is critical to ensuring that research data can be accessed after completion of a study by the tribe and utilized for other purposes (Cross, Fox, Becker-Green, Smith, & Willeto, 2004). Tribes must be able to protect and manage research data that is gathered from their communities to make sure that it is being used in ways that protect their communities and can directly benefit them.

Some tribes may have processes in place regarding research oversight and already do an excellent job of managing research protocols. The purpose of this toolkit is to provide options for tribes interested in developing or updating their research data management practices and policies.

Research Data and Tribal Sovereignty

A research data management plan is one way for tribal nations to exercise sovereignty when they choose to participate in research. A formal data management plan creates an opportunity for data stewardship that aligns with tribal needs and values and establishes the structure necessary for
tribes to harness the power of local data for tribally-driven approaches to improving health.

What Is Data Sovereignty?

Tribes are increasingly interested in ensuring that they exercise their rights as tribal nations with regard to research conducted with their people, lands, and resources. The US Indigenous Data Sovereignty Network at the University of Arizona defines data sovereignty as “the right of a nation to govern the collection, ownership, and application of its own data” (NCAI Policy Research Center, 2018).

Benefits of having a Research Data Management Plan

Having a plan to manage research data accomplish the following actions:

- Improves tribal services and health programs
- Helps tribes determine research priorities and whether or not a proposed project is needed
- Assists in writing grant applications
- Protects data from misuse
- Establishes procedures for how to access data, who can access data, and for what purposes

Does This Toolkit Apply To Every Tribal Community?

NO - there is no one-size-fits-all solution to tribal research data management. Some pieces of this plan may apply to your community, while other pieces are not applicable at all. This toolkit is intended to be used as a guide and a way to get your community thinking and talking about research data management.
# Table of Contents

I. LIFECYCLE OF DATA ................................................................................................................................. 6

II. PRIMARY RESEARCH PROJECT ............................................................................................................ 8
    1. IMPORTANCE OF RESEARCH DATA ................................................................................................. 9
    2. RESEARCH REVIEW ......................................................................................................................... 10
    3. DATA USE AGREEMENT ............................................................................................................... 14

III. SECONDARY RESEARCH PROJECT .................................................................................................... 30

IV. POLICY & PROCEDURE DEVELOPMENT ............................................................................................. 36

V. DEFINITIONS ........................................................................................................................................ 39

V. REFERENCES ......................................................................................................................................... 40
I. LIFECYCLE OF DATA

All research data has its own lifecycle of how it is gathered, used and where it ends up (Figures 1 and 2). This lifecycle will certainly look different for every tribal community. For example, tribal communities with a tribal IRB may not need tribal council approval. The tribal communities should understand their own research data lifecycle as they think about their research data management plan.

Research data comes from two sources: 1) primary data collection, or when the data is first collected in a research project; and 2) secondary data collection, or when the data that was collected from a prior research project is used in another research project. The steps for each of these processes are similar, but may require different data management practices and protections. After all the approvals are completed and the data is collected, the data need to be properly returned to the tribe and securely stored. All of the steps in Figures 1 and 2 are important for tribes to consider defining in their research codes and policies.

LIFECYCLE OF PRIMARY RESEARCH DATA

Figure 1. Example of the lifecycle of primary data - how data can move through a tribal research review process.
Some of the most significant problems that tribes experience with researchers is the unauthorized use of secondary data i.e. data that was collected in the first study that is used for other studies that do not go through the tribal IRB or research approval process. Approval, use, return, and storage of secondary data must go through similar steps as the original research study.

**LIFECYCLE OF SECONDARY RESEARCH DATA**

<table>
<thead>
<tr>
<th>RESEARCH PROPOSAL</th>
<th>RESEARCH REVIEW</th>
<th>APPROVAL</th>
<th>RESEARCH CONDUCTED</th>
<th>APPROVAL</th>
<th>CLOSE-OUT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Analysis Plan is:</strong></td>
<td><strong>Analysis Plan may be Reviewed by One or More of the Following:</strong></td>
<td><strong>Conditions (Outlined in a Data Use Agreement):</strong></td>
<td><strong>Data is:</strong></td>
<td><strong>Before Results are Disseminated:</strong></td>
<td><strong>At this Point, Tribes Must Think About:</strong></td>
</tr>
<tr>
<td>• Clear for new analysis of existing data</td>
<td>• Tribal IRB</td>
<td>• Data only used for approved analysis</td>
<td>• Received by researcher</td>
<td>• Approvals must be obtained according to research review policies</td>
<td><strong>DATA RETURN</strong></td>
</tr>
<tr>
<td>• Specific</td>
<td>• Tribal Committee</td>
<td>• How researcher will get data</td>
<td>• Analyzed according to proposal</td>
<td>• Research review entity must have time to review the presentation, manuscript, report, etc.</td>
<td>• Secure</td>
</tr>
<tr>
<td>• Beneficial to tribe</td>
<td>• Tribal Council</td>
<td>• Return data at close-out</td>
<td>• Tribe owns data</td>
<td></td>
<td><strong>DATA STORAGE</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>DATA RETURN</strong></td>
<td>• Tribal capacity and ability to store</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Filing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Back-up</td>
<td></td>
</tr>
</tbody>
</table>

The review board should ensure the analysis is unique, new, and specific. After this step, results are found and analysis stops here.

**Figure 2. Example of the lifecycle of secondary data - how data can move again through tribal research review process for a new project.**
II. PRIMARY RESEARCH PROJECT
1. IMPORTANCE OF RESEARCH DATA

A significant amount of data has been, and continues to be, collected from AI/AN peoples through research studies. Research data can be any kind of data, not specifically health data. It can include environmental and cultural data. Yet, for many tribal communities, it is challenging to access and manage research data in a way that is meaningful and useful for them. Issues can arise from a variety of situations.

Sometimes, findings are not reported back to tribal communities at all, making it difficult to determine whether any community benefit resulted from participation in the research. Other times, findings are shared with the community but the format does not match the community’s needs or it is not culturally tailored for the community. Some examples of this include the following:

- Using languages or images that are not relevant to the community;
- Providing only summary results (e.g., percentages about the overall population or a particular sub-group), which does not allow for further analysis of the data; or
- Providing only raw data (i.e., data that has not been analyzed) contained in a computer software program that requires additional training or financial expense by the community to be of any use.

Research data can have many important uses:

- It can offer a glimpse into key issues for tribal nations by allowing detection of demographic, health, and social trends;
- It can also help to identify areas for future research and facilitate the establishment of research or health priorities;
- It can be used to improve tribal program services;
- It can be baseline data to apply for additional grant funding or to track indicators over time; and
- It can be used by tribal leaders and stakeholders to make decisions for capacity building.
2. RESEARCH REVIEW

Researchers interested in conducting research with a tribe must go through their research review process. During this process, tribes have the opportunity ask a few questions and to consider several important things when it comes to the research data before approving a proposed project. The research team’s proposal should clearly specify their plans for managing the data that will be collected throughout the project (Table 1). Tribes can find this key information in the research protocol, the initial application, a continuing application, and close-out application.

When developing research review application forms, such as the initial application and close-out applications, tribes could ask specific questions about data and its management. Templates for these forms can be found in the Collaborative Research Center for American Indian Health [CRCAIH] Tribal IRB Toolkit (2015).

Questions regarding research data will vary from tribe-to-tribe but here are a few examples:

- What is the expected study duration through data analysis?
- How will the participant’s privacy be protected throughout the project?
- Which research staff will have access to data that is collected, including participant information, and what is the purpose for their access to this data?
- How will you, the investigator, control who has access to research data?
- Will data be de-identified and coded? If so, when and how will the identifiers be destroyed? Where will the master code be kept?
- What other agencies will have access to any identifiable data?
- How will project documents and data be stored and protected?
- In simple, non-technical jargon, what are the planned analyses for the data?
- Who will be performing the analyses?
- Where will any signed consents or assents be stored?
- How long will you keep data related to the project?
- What are your plans to dispose of the original data?
- What are you plans to return the original data to the tribe?
- As of now, what are your plans to disseminate the results of this project?
• How will project results be returned to participants and community members?
• Have any of the initial plans for data analyses changed? If yes, please explain how and why the change has been made.

Table 1 is an example of a research review timeline and when the research should update the tribe on various parts of the lifecycle of the data that are being collected and analyzed.
Table 1. Example of a research review timeline and when researchers should report to the tribe on the status of the data. As everything in this toolkit, this is simply a suggestion and your tribe may have other factors to consider when determining aspects of research data.

<table>
<thead>
<tr>
<th>DATA ITEMS</th>
<th>Initial Review</th>
<th>Explanation</th>
<th>Continuing Review</th>
<th>Explanation</th>
<th>Close-out</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data Timeline</td>
<td>X</td>
<td>Researcher should indicate their proposed timeline for the project and when the tribe can expect to see data</td>
<td>X</td>
<td>Researcher should update tribe on what stage they are in with data collection, analysis, report writing, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Data Retention</td>
<td>X</td>
<td>This is important so that the researcher and tribe are on the same page about if and how the researcher plans to keep the data to use</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Data Storage</td>
<td>X</td>
<td>The researcher should describe thoroughly how and where data from the project will be stored to ensure it is kept confidential and safe</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Data Analyses</td>
<td>X</td>
<td>Researcher should briefly and simply</td>
<td>X</td>
<td>Researcher should indicate if data</td>
<td>X</td>
<td>Researcher should disclose all analyses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>describe the plans for data analysis</td>
<td>analyses plans have changed since the initial review</td>
<td>that were performed and if they were different than what was proposed at initial and continuing review</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Researchers should have an idea of what dissemination of research results will look like throughout and at the end of the project</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Dissemination</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Researchers should outline plan for the return of the data to the tribe and how they will destroy copies of the data (unless both parties have agreed upon other terms for the researcher’s use and retaining of data)</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Return or Destruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. DATA USE AGREEMENT

Researchers and their teams must plan how they are going to use the data obtained from their study before they begin their project and get approval from the tribe for their plans. Researchers should have an adequate understanding of what the tribe and research review board (if applicable) will require of them in terms of the data they collect throughout their project. Some of these requirements may include:

- How the data can be used or analyzed
- Limits on use of the data
- Prior approval before presentations and publications
- Tribal identification or acknowledgement
- Return or destruction of data
- Agreement with requirements in the tribe’s Data Use Agreement Form

Tribes are encouraged to have every researcher sign a Data Use Agreement Form as a condition of approval of their proposed research study. Data Use Agreements are important since if the researcher and/or their team is noncompliant with the agreed upon terms, tribes are able to take action against the researcher. The researcher should be made aware of what these potential actions could be at the time of signing the agreement. Data Use Agreements are made stronger if they can reference tribal research laws or codes passed by the tribal governing body to regulate research in their community.

A Data Agreement between the researcher and the tribe is a good way to ensure both parties are on the same page and understand the tribe’s expectations when it comes to the research and the data. Tribes should consider requirements in several areas when developing a Data Use Agreement (Figure 3). A Data Use Agreement template is included to provide an example (Example 1, page 21).

It is important to re-emphasize that each tribe will have different needs; therefore, Data Use Agreement forms may look different for each tribe. This toolkit includes examples and meant to serve as a guide for tribes that are developing their research data management. There are a lot of issues to consider in research data management and what works for one tribe may not work for another.
Figure 3. Pieces that should be included in a Data Use Agreement

- Project Purpose
- Dispute Resolution
- Definition of Data
- Data Use
- Data Ownership
- Data Storage
- Data Access
- Data Protection
- Data Collection
- Data Sharing
- Reporting Requirements
- Data Return
EXAMPLE 1. Data Use Agreement Adapted from the University of Washington (University of Washington, 2019):

Data Sharing and Ownership Agreement

for the

<Project Name and Grant name and number>

between the

<Tribe name [Tribe]>

and the

<Institution or Researcher [Institution]>

1. PROJECT PURPOSE AND SCOPE

1.1 Project Purpose: [Tribe] and [Institution] are collaborating to plan, develop, and implement a research study about [topic of research (i.e.) obesity, HIV/AIDS, substance use] within the [reservation or specified boundaries].

1.2 Project Scope: The research project, named [name of study], is [brief description of project including sample size, eligibility criteria, and measures]. This project shall be conducted in accordance with [tribal resolution/code/law] and 45 C.F.R. § 46 et al., along with any and all other applicable federal standards.
governing research on human subjects. This Agreement will be in effect from the date of signature until up to [length of time in years] following the completion of the study.

1.3 **Data Defined:** For the purpose of this Agreement, “Data” shall mean any and all primary source information gathered in the course of the Project that contains “personally-identifiable information”, as that term is used in 45 C.F.R. § 46.101, relating to the Project’s participants. “Data” does not include data sets, data summaries, reports, or other academic publications where personally identifiable information has been removed. The [Tribe’s] cultural traditions, customs, history, and/or stories will never be collected or recorded during this Project, and to the extent any such information is inadvertently recorded in any way, [institution/researcher] agrees to immediately deliver the information to the [Tribe] and destroy any remaining related records in UW’s possession, whether in physical form, electronic form, or otherwise.

2. **DATA MANAGEMENT, SHARING, AND OWNERSHIP**

2.1 **Data Management Guidelines:** The [Tribe] and the [institution] agree to the following binding guidelines for data management, sharing, and ownership:

a) **Data Ownership:** The [Tribe] shall maintain sole ownership and control of the project’s Data. The [institution/researcher] and [funding source, if applicable] are hereby granted a royalty free, nonexclusive right to use the Data to develop and publish reports or other academic publications for educational and research purposes in furtherance of the aims of the Project, provided that the strict confidentiality of all Personally Identifiable Information is maintained at all times and prior approval is obtained from the [name of tribal research review board].
b) **Data Access:** Any entity other than [institution/researcher], [funding source] or the Tribe who wishes to access, view, or use the Project Data must formally request in writing and be granted express permission by the [tribal research review board/tribal council/committee].

c) **Data Collection:** Before any Data is collected, [researcher name] must have an approved IRB application and approval by the [Tribal IRB or other review].

d) **Data Protection:** [Researcher name] is responsible for data collection and analysis, and for keeping all data containing personally identifiable information and their sources confidential. By signing this agreement, [researcher name] recognizes and accepts responsibility to ensure appropriate investigator conduct in relation to the data. [Researcher name] shall remove all Personally Identifying Information from any and all Data collected as a result of the Project in accordance with applicable federal, state, tribal, and institutional laws and regulations.

e) **Data Storage:** For the duration of the Project, [researcher name] will house all physical and electronic Data in secured, locked physical and electronic systems, including locked file cabinets and secure, password-protected databases on secure servers within the [location] offices. [Researcher] will also maintain and keep the de-identified Project Data to the limited extent necessary to complete any study aims, reports, and dissemination activities authorized by the IRB. After termination of the Project and upon completion of the time allotted by the IRB to fulfill Project-approved publication or dissemination activities, all data retained by [researcher name] must be provided to the [Tribe], or destroyed at the [Tribe] express written direction.
f) **Data Use**: In connection with the project data, the parties agree to establish a Dissemination Plan to utilize and disseminate the Data, to both academic audiences and the [Tribe] community. No project results will be disseminated without approval by the [Tribal IRB]. If this is an initial project, any secondary analyses that could be conducted must be approved with a new IRB application and review.

g) **Reporting Requirements**: Upon completion of the project, [researcher name] will provide the [Tribe IRB] with a community-specific report and an aggregate report that includes findings from the project.

2.2 **Ownership of Reports/Publications**: The [Tribe] and [researcher name] shall jointly maintain ownership of all Project reviewed and approved reports and publications in which [Tribe] and [Institution] researchers are co-authors.

3. **DISPUTE RESOLUTION**

*In case of a dispute or an instance that the researcher or tribe breaks this agreement, the tribe should think about what steps must be taken in and what will be needed to enforce them. It is suggested to consult with tribal lawyers to ensure appropriate language is used in cases where tribal laws may need to be enforced.*
4. DATA RETURN

Tribes and researchers must agree before project initiation on the data management and return process. The most important agreement must be that any research data obtained from a project is owned by the tribe and must be returned to the tribe after the proposed project is completed. Researchers must return the data to the tribe according to the tribe’s data management and return process. The return of the data must be documented by the tribe and the researcher. Any copies of the data in possession by the researcher must be destroyed after they have returned the primary data to the tribe unless there is an agreement made between the tribe and researcher. They cannot use that data for another research project or analysis unless they get approval from the tribe.

The following flowchart is an example of how the data return process might look for a tribe:
Tribes can develop forms or other methods of documentation that the researcher must complete when they return the data. The Oglala Sioux Tribe Research Review Board Data Return Form is included below in Example 2.

**Example 2. DATA RETURN FORM** (pages 24-26). Adapted from the Oglala Sioux Tribe Research Review Board:

<table>
<thead>
<tr>
<th>Name of Principal Investigator:</th>
<th>Phone:</th>
<th>Email:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mailing Address:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of Contact Person:</th>
<th>Phone:</th>
<th>E-mail:</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Co-PI/s ☐ Other:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Institution/Organization:</th>
<th>Sponsor (if funded):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project Title:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project Start Date:</th>
<th>Project End Date:</th>
<th>Name of Tribal Research Office Approval Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other IRB Approvals:</th>
<th>Keywords:</th>
</tr>
</thead>
</table>
Please list information on all journal articles, poster presentation (poster & oral), in the following table. If you need more space, please include a separate document with table.

<table>
<thead>
<tr>
<th>Title</th>
<th>Type of publication</th>
<th>Date</th>
<th>Location</th>
<th>Conference/Journal name</th>
<th>Presenter/Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Possible future uses of the data – please identify ways in which this data could contribute to future analysis/studies (specify research topics, keywords, or questions):

Along with this form please include on a single flash drive containing the following items:

- Data folder containing two subfolders
  - Data folder – Containing de-identified data in universal format (.txt or .csv) and any software specific formats
  - Codebook and Instrument folder – Containing a readme.txt and codebook.pdf file to fully explain all of the variables in the raw set, as well as a blank copy of the survey instrument that was used
- Grant and relevant IRB forms – a copy of all relevant IRB closure forms
- If applicable, Newsletters and other forms of dissemination of results

ASSURANCES

I certify that all data from the study conducted within *(ex: the exterior boundaries of the [name] Reservation)* is included per the approval letter from the IRB. If not, the IRB has the rights to request missing data.
Please sign below where indicated and return this form and the flash drive to

Certified Mail

FedEx/UPS

Also email form to ________________________________ to facilitate processing.

____________________________________  __________________
Investigator’s Signature                  Date
5. DATA STORAGE

Good data stewardship requires that data is properly maintained and stored once it is returned to the tribe. Beyond the simple act of holding data, tribal data stewardship acknowledges that information collected through research studies with tribal communities may cover a range of topics from low-risk to highly sensitive individual and cultural information, all of which is valuable by virtue of the fact that is has been provided by consenting tribal citizens and has the potential to impact tribal wellbeing. This data must be securely held and managed in a way that preserves it, protects the research subjects and the tribe, and allows for efficient access for future use.

If the tribe determines that it will collect research data, there should be a secure location to store the data. There may be an Information Technology (IT) department available for assistance with storing electronic data or it can be done by an individual with a minimal infrastructure.

Data Access and Security

In order to secure research data and prevent it from being altered or corrupted, access to the data should be limited. Allowing only certain people to access the data will hopefully prevent accidents like erasing or moving files, changing data or filenames, or improper usage of data. Identifying the individuals that can access data files is essential. While access should be limited, there should always be more than one individual that has access to files. It is necessary to have more than one person with control over the data storage just in case the original holder loses accessibility or leaves their job.

Data should be password protected, with access to the password limited to the individuals that have access to the data. Data could be rendered unusable if it is stored on a computer that is infected with a virus or other malicious software, otherwise known as malware. Anti-virus and anti-malware software should also be taken into consideration to prevent against potential cyber-attacks. Other possible protections to think about include a firewall and data encryption scheme.

Location

Storage location is also important to keep in mind when working with data from research studies. Some types of data files, such as videos, audio recordings, and images take large amounts of storage space.

There are many options available for storage of data contained within electronic or computer files. Three possible options include storage on a local
device, cloud storage, and file servers storage. Each option has advantages and disadvantages and each tribe will have to determine which option is best for its community (Table 2).

**Data Back-Up**

As mentioned above, data should be backed up on a regular schedule regardless of where it is stored using a well-documented process. Backing up data means making at least one or more additional secure copies of the data and storing them in separate locations in case something bad happens to the original data files. The simplest example of this would be copying the entire contents of a disk drive to a separate disk drive. It is typically recommended that there are at least two backup copies of the data plus the original. It is also a good idea to store at least one of the backup copies of the data in a separate physical location so that it is not harmed if there is physical damage, like a fire or flood, in the location of the original.

The data backup process needs to include a specific timeline that is followed. Data could be backed up daily, weekly, or monthly. The more frequently that data is backed up, the more likely it is to be recovered in its currently form if needed. However, data backup takes time so it is recommended to make a plan and stick to it. Unfortunately, one missed backup can lead to data loss and project delay.

Documentation of the back up procedure is required when backing up data. It is also recommended that the process be tested to ensure that data can be successfully retrieved when the backup is needed. Testing a data backup procedure could be as simple as periodically utilizing one of the backup copies to attempt to access data files, verifying that all files are able to be located and the data can be utilized.

The following table (Table 2) highlights advantages and disadvantages of using cloud storage, a dedicated server, and local storage for storage and backups.
Table 2. Advantages and disadvantages of different storage types.

<table>
<thead>
<tr>
<th>STORAGE TYPE</th>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOUD</td>
<td>-Easy to use</td>
<td>-Data is not local</td>
</tr>
<tr>
<td></td>
<td>-Backup/Recovery is covered by service provider</td>
<td>-Monthly service fee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-CRCAIH Tribal Partners have brought up security concerns with this type of storage</td>
</tr>
<tr>
<td>SERVER</td>
<td>-Locally managed by trusted professionals</td>
<td>-IT department needs to support/maintain servers</td>
</tr>
<tr>
<td></td>
<td>-Backup/Responsibility is responsibility of Tribe’s IT department</td>
<td>-Equipment is expensive to purchase and maintain</td>
</tr>
<tr>
<td>LOCAL STORAGE (HARD DRIVES)</td>
<td>-Data is very accessible</td>
<td>-Discs can be corrupted or damaged</td>
</tr>
<tr>
<td></td>
<td>-Data is local</td>
<td>-Backup takes time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Limited lifespan</td>
</tr>
</tbody>
</table>
**File/Folder Naming**

When data is going to be stored, it should be organized in a systematic fashion. An example of a systematic organization scheme is the Dewey Decimal System that is used to organize books in a library.

A clear method of organizing the data files from research studies will allow the tribe to locate and utilize the data in an efficient manner. Personnel responsible for storage and retrieval of data may change over time, and one person’s organization scheme may not be useful to someone else. If there is a consistent naming scheme that is used it will be much easier to locate data when it is requested.

Begin by setting up a file directory (folder) structure that consistently organizes files from separate studies. Using descriptive terms in folder names can help identify the contents of the folder.

A clearly defined file naming scheme for research data storage should also be implemented and documented. File names should describe what is contained in the file. For example, the filename “document1.txt” does not give any indication of what is in the document but “data_management_plan.txt” does describe the file contents. Each file name also should contain a file extension at the end. In the examples listed “.txt” is the file extension. A file extension informs a computer’s operating system which program can be used to open the file. Below are some recommendations for file naming (Organizing Your Data):

- Try to keep the file names relatively short (under 25 characters)
- Use the underscore character ( _ ) rather than spaces in file names
- Avoid using symbols in file names
- There should only be one period ( . ) in a filename (part of the file extension)
6. USE BY TRIBE

When the primary data has been collected as a result of the research study and the results are ready to be shared with stakeholders and community members, researchers must follow through with their dissemination plan. The data and knowledge gathered should be used to strengthen or benefit the community in which the project was conducted. Dissemination must be considered from the beginning of the project to the end, with continuous consideration throughout.

If a tribal research review board is available, this group would be where the tribe can monitor a researcher’s dissemination plan (Example 3 on page 29). This dissemination plan should be developed collaboratively with the researcher and tribal stakeholders, whether that is a specific committee or the community partners assisting with the project. The researchers should understand the time commitment this may take and be able to commit staff time to these efforts. The dissemination plan is an incredibly important piece of a research project because it allows data obtained to be translated into real life situations.

If there are the capabilities, evaluation of the dissemination efforts to determine the outcomes would be beneficial. This could be done holding focus groups, interviews, and surveys, evaluating the materials and how they have been disseminated.

When data is returned to a tribe, it will be able to be accessed by tribal entities who need it for:

- Developing policies
- Writing grants
- Reports
- Secondary research
- Community education

Questions to consider when looking at a researcher’s dissemination plan:

- Who do they know in the community?
- What groups or programs does the data pertain to?
- What places have they identified to hang up flyers?
- Will they post on social media (Twitter, Facebook, Instagram)?
- Are they going to work with community partners to develop culturally tailored materials?
Example 3. DISSEMINATION PLAN TEMPLATE (page 36). Adapted from the University of Regina, Community Research Unit:

**DISSEMINATION PLAN WORKSHEET**

<table>
<thead>
<tr>
<th>PROJECT DETAILS</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESEARCH PROJECT</td>
</tr>
<tr>
<td>RESEARCH PARTNERS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISSEMINATION PLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>TARGET AUDIENCE(S)</td>
</tr>
<tr>
<td>OBJECTIVES</td>
</tr>
<tr>
<td>DISSEMINATION METHOD(S)</td>
</tr>
<tr>
<td>LOCATIONS (OFFICES, STORES, SOCIAL MEDIA ACCOUNTS, ETC.)</td>
</tr>
<tr>
<td>RESOURCES &amp; FUNDING NEEDED</td>
</tr>
<tr>
<td>RELEVANT TRIBAL PROGRAMS</td>
</tr>
<tr>
<td>TIMELINE</td>
</tr>
<tr>
<td>ASSIGNMENT OF TASKS</td>
</tr>
</tbody>
</table>
III. SECONDARY RESEARCH PROJECT
IDENTIFYING USEFUL SECONDARY DATA ANALYSES

The primary research data collected in a research project can also have potential use in another “secondary” research project. Research data already collected could be useful for tribal entities for items such as reporting purposes, additional research projects, or as a baseline for applying for additional grant funding. Besides specific research data, aggregate data from reports, publications, and publicly available databases can be readily adapted to meet tribal needs. Many of these resources are available online.

Tribes must know what types of data are available before that data can be used in another project. The Regulatory Knowledge Core of CRCAIH has helped tribal partners catalog previous research to provide a listing of studies that have taken place. This method of cataloging studies allows the tribe to identify research studies that have been completed, along with information about those studies including what type of data is available for future studies.

Tribes must also keep contact information on the original researcher for each project in case another researcher wants to use the data and has questions. The original researcher is familiar with the data and can help to identify whether data from their study will be useful for the potential secondary use. They also may have other potential resources that could assist with the proposed analysis.

REGULATORY CONCERNS

Using research data to explore topics or questions that were not included in the original research project can lead to several questions.

- Can this data be used to explore additional topics and questions?
- What are the terms of the consent form from the original research project? Is additional consent required?
- Are there additional regulatory approvals that are needed to use the data?

In addition to knowing the type of data that will be used, it is also important to know how data is going to be used for secondary analysis. How the data will be used often determines whether additional IRB reviews and approvals are needed.

Data from completed research studies can potentially be repurposed to provide additional information for tribes. However, before this data is analyzed for another purpose, the intent of the original research study must be thoroughly examined. Additional use of individual level data needs to be
consistent with the original consent obtained from the participant. If it is not consistent with the original consent, then the secondary analysis cannot take place without approval of the study both by review boards and the individual participants.

Approval for secondary analysis of individual level data also depends on whether or not participants can be identified, how data is shared, who owns the data, as well as how research policy is enacted. If personal information that could be used to connect data with a specific person has been removed from the data set, then it is referred to as de-identified data. This impacts whether IRB review is needed to ensure protection of human subjects. However, tribal review is likely even if IRB review is not required. A comparison of the potential requirements depending on whether the data contains identifiers is include below:

- **If data is de-identified:**
  - No longer meets the definition of human subjects research from the Office for Human Research Protections
  - May not need approval from some review boards
  - May need tribal community approval

- **If data is identified:**
  - May need additional approvals for use – would likely need to go back to all original review boards
  - May need tribal community approval
  - May need additional consent from the study participants

The recently updated Common Rule, which includes regulatory requirements for researchers conducting research with certain federal funding, includes some important changes in the handling of informed consent and research review. If the consent form for the original project was signed by the participant and they agreed to “broad consent”, their identifiable data can be used by other researchers in any future studies. However, some tribes and institutions are deciding to not approve research studies with broad consent in their informed consent forms (NCAI Policy Research Center, 2019).

Tribes should also reaffirm through their research laws or codes what types of review are required for use of both primary and secondary data. Even though data without identifiers might not require IRB review to assess human subjects protections, tribes can still require tribal review of all research studies, whether they use primary or secondary data (NCAI Policy Research Center, 2019).
USES FOR PRIOR RESEARCH DATA

Researchers often use primary data collected in a past research study in new research studies that involve secondary data analyzes to answer new or additional research questions. Sharing data among researchers is considered beneficial since it is more efficient and can result in findings that can help improve health or reduce disparities much sooner than having to recreate each study.

Tribes may have some important uses for secondary data analyzes. Secondary data can assist in development of community assessments and grant proposals, including data aggregated or combined from multiple datasets. When multiple data sets are compiled over time, it allows the tribe to create a snapshot of what is happening in health research. If there are collections of data that track certain factors over a long period of time, an aggregate data set could be used to examine trends over time.

Community assessments involve community level data and can include a number of different topic areas such as health, environment, housing, etc. The most common way to collect data for a community assessment involves a survey of individuals within the community. Previous research data can help inform the design of the study and highlight important issues that should be addressed by items on the survey. One example of this is developing questions for a community assessment by using data from a prior study that performed focus groups on health care availability.

Additionally, aggregate data can also be used as preliminary data for grant proposals. Community level data can identify areas of importance for future research and help the grant writer present a case that a research study is needed in a particular area in a particular community. Having local data helps create a strong grant proposal.

A good data management plan can allow for accurate aggregation or combining of datasets from prior research studies.
SOFTWARE FOR DATA ANALYSIS

Secondary analysis involves taking previously gathered data, asking a new research question, and analyzing the data for new results. Research data is usually stored in databases with multiple records and variables using excel spreadsheets or, more commonly, research/statistical software.

In order to analyze the data to discover patterns or calculate statistics, computer software is generally used. There are many different software packages available for data analysis. Choosing a software package that meets the data analysis needs of the tribe is dependent on many factors. One of these factors is price. Statistical software packages can range in price from free to thousands of dollars per license. Another factor is the ease of use. This can depend on the user’s level of comfort with the software program, the availability of training and support, and much more.

The CRCAIH Methodology Core has several trainings on using computer software for data analysis available on the CRCAIH website. Methodology Core members also have expertise with many statistical software packages and can help provide guidance in choosing the package that would be a good fit.
IV. RESEARCH DATA POLICIES & PROCEDURES

*This section introduces a proposed template for a tribe’s Policies & Procedures in handling research data. These are simply factors that would be helpful when thinking of research data return and what steps may need to be taken by your tribe. The items listed are meant to help you explore your tribe’s needs and capabilities when developing a Policy & Procedure around research data and managing it.
RESEARCH DATA MANAGEMENT POLICY & PROCEDURE DEVELOPMENT

A. Introduction

- Purpose and Applicability of Research Data Policies and Procedures
  - Broad description of document contents
  - Who must abide by these policies and procedures? (e.g. researchers, tribal IRB members)
  - Relationship of policies and procedures to tribal law and research review policies
  - Definitions

B. Overview of Research Data

- Policies regarding Research Data
- Roles and Responsibilities
  - List who is responsible for gathering research data from completed studies, verifying that requested data has been returned, storing data according to specifications, and backing-up the data

C. Data Return

Researchers are responsible for data collection and analysis for their research project. They shall carefully keep any identifiable information confidential and remove any personal identifiers according to regulations. The researcher will store all physical and electronic data in secure locations that can only be accessed by identified personnel. Upon termination or completion of the study, or more frequently as directed by the Tribal IRB, the principal investigator is responsible to return all physical and/or electronic data to the Research Data Manager with the Tribal Research Office.

- The following are examples of the different types of data the Tribal Review Board may require you return.
  1. Physical-understanding that any physical data shall be returned using a secure option through a reputable delivery service.
  2. All electronic data shall be returned in duplicate on identical password protected flash drives that are mailed securely through a reputable delivery service.

- Standards for Data and Metadata
1. Any physical data shall be returned using a secure option through a reputable delivery service. The data that shall be included on the flash drive includes:

2. A Directory (folder) named DATA which contains de-identified data in a universal format. Software specific format can also be included but is not required.

3. A Directory named CODEBOOK AND INSTRUMENT that contains a codebook or data dictionary in a universal format (.PDF and/or .TXT) as well as an original copy of the instrument that was used to collect the data.

4. A Directory named FORMS that contains other relevant IRB forms related to the project.

5. A Directory named RESULTS that contains any presentations or publications that were a direct result of the project.

iii. How it will be returned

1. Outline the specific requirements of returning the data

D. Presentations & Publications
   • Specify what, if any, approval are needed prior to a researcher publishing or presenting research data
   • Outline tribal identification processes

E. Data Storage
   • Where will the data be stored?
   • Identify procedures for data back-up
   • Who will have access to and manage this data?
F. Data Sharing

- How will your tribe handle requests for secondary data analyses or use of data for grant writing or program improvement/development?
- Who will be allowed to use this data?
- How will your tribe share this data with the interested party?
- How long will it take for the interested party to receive data once it is requested?

Data Agreement
- Specify if interested party and tribe must enter into an agreement prior to the sharing of this data

G. Secondary Data Analysis

- In order to utilize prior research data, a request letter should be sent to the Institutional Review Board Chair (IRB) as well as the Research Data Manager. The letter should request specifically what data is requested to be used, the purpose for the request, and other relevant information. The IRB Chair and the Research Data Manager will then determine whether to allow the data to be accessed and utilized, as well as if any additional regulatory actions are required to use the data.
V. DEFINITIONS

**De-identified Data**: Data that has had all identifying information removed so that the data cannot be linked back to the person who provided it.

**Raw Data**: When the data is first collected, this is called the primary or raw data. This data has not had anything done to it; it is whatever was collected initially. It has not been subjected to processing or any other analysis.

**Primary Research Data**: Data that is collected by a researcher and their team using first-hand techniques, such as interviews, surveys, focus groups, or other experiment.

**Secondary Research Data**: This is Primary Research Data that has already been collected but is being used for a new project, such as using a dataset from a research project to explore a new hypothesis through different analyses techniques.

**Research Data**: Recorded factual material commonly accepted in the scientific community as necessary to document and support research findings (NIH Data Sharing Policy and Implementation Guide, 2003)

**Local Storage**: Files stored on local devices such as the hard drive of a personal computer, external hard drive, flash drive, or a CD/DVD. Storing on a local device introduces risk since these drives are likely to have a limited lifespan and can be easily damaged. Secure back up of local storage is required.

**Cloud Storage**: Storage requires a secure internet connection to upload, download, and manage data files.

**File Server Storage**: A file server dedicated in local system that is usually accessed via a network by multiple computers.

**Data Backup**: Making multiple copies of the data and storing them in separate locations. One example of this is copying the entire contents of a disk drive to a separate disk drive.

**Community Assessments**: Community level data, which can include a number of different topic areas such as health, environment, and housing.

**Dissemination**: Concerning research, dissemination refers to communicating the results back to stakeholders, the scientific community, and community members. Dissemination can come in many forms and is not only meant to refer to journal publications and poster presentations; it may also refer to community presentations, infographics, and community education materials, etc.
VI. REFERENCES


University of Washington. (2019). Data
Wopida Tanka, Wopila Tanka, and Miigwech

Pictured: Melissa Buffalo (Meskwaki)
Taken by: Michaela Seiber (Sisseton-Wahpeton Oyate)