

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

MCC Evaluation Microdata
Data Package

Instructions

This template is informed by MCC's Evaluation Microdata Documentation and De-Identification Guidelines. In addition to reviewing these Guidelines, MCC contractors responsible for preparation and documentation of evaluation-related microdata for public and/or restricted-access use should be familiar with the following US government guidelines for data de-identification and re-identification:

- [NIST 2015](#)
- [NIST 2016](#)

MCC, the evaluator, and stakeholders should consider the following multi-stage process for data review and release:

1. Evaluator and M&E PM should agree on expected DRB review date as early as possible to confirm. This should be scheduled at least one month before Evaluator's contract expires.
2. Evaluator should submit full package to M&E PM. The package includes:
 - One completed Section 1 of the DRB Data Package Worksheet for ALL data components (i.e. individual, household, and community data for one survey round are three data components with different risks)
 - One completed Section 2 & 3 for EACH data component
 - Datasets and code package(s)
 - Informed consent(s)
 - Questionnaire(s)
 - Most recent Metadata file (for [Evaluation Catalog](#) entry)
3. M&E PM should review Metadata and DRB Data Package Worksheet for clarity and completeness. This may require one round of revision based on the M&E PM requests for clarity and completeness.
4. Evaluator should submit full package to M&E PM. M&E PM and the M&E DRB members should establish a first-round review and feedback to the Evaluator on the proposed data de-identification process. This may require a second round of revision to the package.
5. Evaluator should submit full package to M&E PM for the confirmed MCC DRB review date at least 2 weeks prior to confirmed DRB review date.
6. If any feedback/revisions are required following MCC DRB review, Evaluator should revise and resubmit full package to M&E PM with documented responses to MCC DRB feedback to ensure timely virtual review and clearance of the full package. All final de-identification efforts and their impact on verification of analysis should be documented in the evaluator's Transparency Statement available on the Evaluation Catalog.

All **red font text** are instructions in the Worksheet and must be replaced with standard black font with the contractor's response.

Unless otherwise agreed with MCC, the final document will be made public to complement/underlie the contractor's Transparency Statement to document the data preparation and de-identification

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

process required for the public and/or restricted-access microdata and any impact on the data for verifying evaluation analysis and broader data usability.

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

Section 1: Cover Sheet

Overview of Data Package

(Instructions: Include a paragraph summarizing each data package component included in the package. For example, if the package includes household, individual, and community level data sets, please include a paragraph summarizing each of these three components, including information on the content and timing of the data collection.)

This data package includes the following components:

Data Package Component 1: Patron Survey

This package contains datasets for two surveys of individuals who are users of Namibia's three RSRC's. The surveys were conducted in July-September 2017 and July-September 2018. "Users" are defined as those who have visited the RSRC two or more times in the previous 6 months. Only individuals 15 years old or above were sampled. Instrument content includes questions around basic demographics, usage activities and frequencies, and satisfaction/perceived outcomes.

Data Package Component 2: Panel Survey

This package contains datasets for two surveys of Namibia's three RSRC's, as part of the panel study. Survey participants are individuals who self-identify into one of two primary occupation groups 1) high school student or 2) "business patron" (i.e. business owner/entrepreneur, wage earner, or job seeker). "Users" are defined as those who have visited the RSRC two or more times in the previous 6 months before recruitment. The Panel Survey population is a subset of those sampled in the Patron Survey. Instrument content includes questions around basic demographics, RSRC usage activities and frequencies, reasons for use/non-use, and perceived impact of the RSRC. Data collection activities for the datasets included in this package were repeated in December 2017 and November 2018.

Data Package Component 3: Panel Interviews

This package contains datasets for two interviews of Namibia's three RSRC's, as part of the panel study. Participants are individuals who are users of Namibia's three RSRC's and self-identify into one of two primary occupation groups 1) high school student or 2) "business patron" (i.e. business owner/entrepreneur, wage earner, or job seeker). "Users" are defined as those who have visited the RSRC two or more times in the previous 6 months before recruitment. The Panel Interview population is a subset of those sampled in the Patron Survey and is the same as those that participated in the panel survey. Instrument content includes questions around basic demographics, RSRC usage activities and frequencies, reasons for use/non-use, and perceived impact of the RSRC. Data collection activities for the datasets included in this package were repeated in August/September 2017 and August/September 2018.

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

Data Package Component 4: Qualitative Methods

Commented [GAA(1): Add note that these data will not be submitted to MCC and why, e.g., IRB did not support this. The worksheet should also mention the site observation dataset and why it won't be shared.

Data Package Folder Contents

(Instructions: Please list the Data Package Component File Name, and then include the File Names of each of the corresponding required documents [Metadata, Worksheet, Informed Consent, Questionnaire, Other docs]. Only one de-identification worksheet per survey is requested unless discussed.)

Table 1: Data Package Components

Data Package				
Component	Worksheet	Informed Consent	Questionnaire	Other Documents
Component_1_Patron_Survey	DRB_Data_Package_Worksheet_Quant_Final.docx	<ul style="list-style-type: none"> Research_Subject_Information_Sheet_PatronSurvey TEST_0216_signed.pdf Research_Subject_Consent_Script_PatronSurvey_062717.docx Research_Subject_Minor_Assent_Script_PatronSurvey_062817.docx 	<ul style="list-style-type: none"> Instrument_PatronSurvey1.docx Instrument_PatronSurvey2.docx 	<ul style="list-style-type: none"> Data_PatronSurvey1.csv Data_PatronSurvey2.csv RMD_PatronSurvey1.rmd RMD_PatronSurvey2.rmd Codebook_PatronSurvey1.pdf Codebook_PatronSurvey2.pdf
Component_2_Panel_Survey	DRB_Data_Package_Worksheet_Quant_Final.docx	<ul style="list-style-type: none"> Research_Subject_Information_Sheet_Panel_Student TEST_0216_signed.pdf Research_Subject_Information_Sheet_Panel_Biz TEST_0216_signed.pdf Research_Subject_Consent_Script_Panel_Participants.docx Research_Subject_Minor_Assent_Script_Panel_Participants.docx 	<ul style="list-style-type: none"> Instrument_StudentPanel_Survey.docx Instrument_BusinessPanel_Survey.docx 	<ul style="list-style-type: none"> Data_Student_Panel_Survey1.csv Data_Student_Panel_Survey3.csv Data_Business_Panel_Survey1.csv Data_Business_Panel_Survey3.csv RMD_StudentPanel_Survey_1VS3.Rmd RMD_BusinessPanel_Survey_1VS3.Rmd Codebook_StudentPanel_Survey1.pdf

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

				<ul style="list-style-type: none"> • Codebook_StudentPanel_Survey3.pdf • Codebook_BusinessPanel_Survey1.pdf • Codebook_BusinessPanel_Survey3.pdf
Component_3_Panel_Interviews	DRB_Data_Package_Worksheet_Quant_Final.docx	<ul style="list-style-type: none"> • Research_Subject_Information_Sheet_Panel_Student_TEST_0216_signed • Research_Subject_Information_Sheet_Panel_Biz_TEST_0216_signed.pdf • Research_Subject_Consent_Script_Panel_Participants.docx • Research_Subject_Minor_Assent_Script_Panel_Participants.docx 	<ul style="list-style-type: none"> • Instrument_StudentPanel_Interview.docx • Instrument_BusinessPanel_Interview.docx 	<ul style="list-style-type: none"> • Data_Student_Panel_Interview_Initial.csv • Data_Student_Panel_Interview_Final.csv • Data_Business_Panel_Interview_Initial.csv • Data_Business_Panel_Interview_Final.csv • RMD_StudentBusinessPanel_Interview_1vs2.Rmd • Codebook_StudentPanel_Interview_Initial.pdf • Codebook_StudentPanel_Interview_Final.pdf • Codebook_BusinessPanel_Interview_Initial.pdf • Codebook_BusinessPanel_Interview_Final.pdf

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

Section 2:

Data Component #1 (Patron Survey) Preparation Overview

	Response		Discussion/Explanation
Data + Code Completeness	Complete		<i>To be considered Complete: The available data must allow new users to replicate evaluator analysis to the extent allowable by providing the full data set + analysis code. The constructed variables may also be included in a dataset, but if the dataset+code produces those variables, it is not necessary.</i>
	Incomplete	The data (as .csv) and R code (as .rmd) are complete and are contained in this package. Also created to support reproduction are a code book.	<i>To be considered Incomplete: The available data only provides a sub-section of data as produced by the survey and/or the constructed variables only. Incomplete data files are limited in terms of full verification of analysis and/or broad usability of data and must be justified.</i>
Data Round(s):	Baseline only	Combination of rounds. The package contains data from the first and final rounds of the patron survey (2017 and 2018).	<i>MCC is willing to trade-off broad use of individual rounds for more consistent de-identification protocols across rounds of data. Therefore, unless there is specific demand for the baseline/interim only data, or contractual requirements, MCC prefers contractors to prepare all data rounds in one package.</i>
	Interim only		<i>If one stage only – please (i) confirm demand and/or contractual justification and (ii) discuss how preparation and release of this data as presented to the DRB may affect future data round releases.</i>
	Endline only		
	Combination of rounds		<i>If combination, please discuss if this file replaces any previously published datasets.</i>
Informed Consent and IRB	High restriction	Given the nature of questions in the patron survey and the study overall, low restriction was required. As such and	<i>MCC assumes DIRECT identifiers are always removed from any public-use file. With this assumption: Please refer to the informed consent statement – does it require: High restriction: access</i>

Formatted: Font: Bold

Formatted: Font: Bold

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

	Medium restriction	following expectations laid out in the consent script, only direct personal identifiers were removed from the dataset. The informed consent script states "TASCHA and MCC will remove any identifying information from the transcripts and notes – such as names, dates, and specific locations." In addition, given the general nature of the questions and large population geographies and sizes that surround the RSRCs, we determined a very low risk of re-identification from indirect identifiers.		to data that includes indirect identifiers is limited to the contractor only; Medium restriction: access to data that includes indirect identifiers is limited to the contractor and qualified researchers, including MCC; Low restriction: data with indirect identifiers may be made public.
	Low restriction			Please discuss how the promises of confidentiality in the informed consent informed de-identification efforts. Please include any additional guidance provided by the IRB as applicable.
Geographic Identifiers	Highest (Region)	Oshakati (pop. 182K) Omaheke (pop. 71K) Ohangwena (pop. 245K)	The regions are so large and populated that individual identification through this variable is considered nearly impossible.	Please provide justification on the identification/de-identification/complete removal of specific geographic regions. De-identifying at a higher geographic level may support privacy protection, but it may also reduce data usability. Please provide justification for recommendation.
	--(i.e. District)	NA – Not collected		
	--(i.e. State)	NA – Not collected		
	--(i.e. Village)	NA – Not collected		
	Lowest --(i.e. Census Blocks)	NA – Not collected		
Knowledge of Treatment	High risk	Knowledge of the survey taking place at the regional library is assumed to provide little to no re-identification risks. Given the number of patrons that use		In some cases, general knowledge of treatment areas and/or inclusion of a treatment variable can significantly increase re-identification risk depending on the population affected. Please provide assessment of this re-identification risk
	Medium risk			

Formatted: Font: Bold

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

	Low risk	each library every year and the overall non-personal nature of survey questions, we deem re-identification a very low risk.	<i>and recommendation if considered high/medium risk.</i>
Publication Type	Public-use only		
	Restricted-use only	The package contains public-use data only. Codebooks, and analysis code files will allow for verification of evaluation results.	<i>Please state for this data package: will there be public-use data only, restricted-use data only, or both and provide justification as this relates to enabling verification of evaluation results and/or broad usability of the data.</i>
	Both		

Formatted: Font: Bold

Formatted: Font: Bold

Data Component #2 (Panel Survey) Preparation Overview

	Response		Discussion/Explanation
Data + Code Completeness	Complete		<i>To be considered Complete: The available data must allow new users to replicate evaluator analysis to the extent allowable by providing the full data set + analysis code. The constructed variables may also be included in a dataset, but if the dataset+code produces those variables, it is not necessary.</i>
	Incomplete	The data (as .csv) and R code (as .rmd) are complete and are contained in this package. Also created to support reproduction are a code book.	<i>To be considered Incomplete: The available data only provides a sub-section of data as produced by the survey and/or the constructed variables only. Incomplete data files are limited in terms of full verification of analysis and/or broad usability of data and must be justified.</i>
Data Round(s):	Baseline only	Combination of rounds. Contained in this package are rounds 1 and 3 of data collection, which occurred in December 2017 and November 2018, respectively. Round 2 was not used in the final	<i>MCC is willing to trade-off broad use of individual rounds for more consistent de-identification protocols across rounds of data. Therefore, unless there is specific demand for the baseline/interim only data, or contractual requirements, MCC prefers contractors to prepare all data rounds in one package.</i>
	Interim only		

Formatted: Font: Bold

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

	Endline only	analysis and, as such, not included in this package.		If one stage only – please (i) confirm demand and/or contractual justification and (ii) discuss how preparation and release of this data as presented to the DRB may affect future data round releases.
	Combination of rounds			
Informed Consent and IRB	High restriction	Given the nature of questions in the panel survey and the study overall, low restriction was required. As such and following expectations laid out in the consent script, only direct personal identifiers were removed from the dataset. The informed consent script states “TASCHA and MCC will remove any identifying information from the transcripts and notes – such as names, dates, and specific locations.” In addition, given the general nature of the questions and large population geographies and sizes that surround the RSRCs, we determined a very low risk of de-identification from indirect identifiers.		MCC assumes DIRECT identifiers are always removed from any public-use file. With this assumption: Please refer to the informed consent statement – does it require: High restriction: access to data that includes indirect identifiers is limited to the contractor only; Medium restriction: access to data that includes indirect identifiers is limited to the contractor and qualified researchers, including MCC; Low restriction: data with indirect identifiers may be made public.
	Medium restriction			
	Low restriction			
				Please discuss how the promises of confidentiality in the informed consent informed de-identification efforts. Please include any additional guidance provided by the IRB as applicable.
Geographic Identifiers	Highest (Region)	Oshakati (pop. 182K) Omaheke (pop. 71K) Ohangwena (pop. 245K)	The regions are so large and populated that individual identification through this variable is considered nearly impossible.	Please provide justification on the identification/de-identification/complete removal of specific geographic regions. De-identifying at a higher geographic level may support privacy protection, but it may also reduce data usability. Please provide justification for recommendation.
	--(i.e. District)	NA – Not collected		
	--(i.e. State)	NA – Not collected		

Commented [GAA(2)]: Should we obtain this as a separate file since we paid for it?

Commented [SJ(3R2)]: Yes.

Formatted: Font: Bold

Formatted: Font: Bold

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

	--(i.e. Village)	NA – Not collected		
	Lowest --(i.e. Census Blocks)	NA – Not collected		
Knowledge of Treatment	High risk	Knowledge of the survey taking place at the regional library is assumed to provide little to no re-identification risks. Given the number of patrons that use each library every year and the overall non-personal nature of survey questions, we deem re-identification a very low risk.		<i>In some cases, general knowledge of treatment areas and/or inclusion of a treatment variable can significantly increase re-identification risk depending on the population affected. Please provide assessment of this re-identification risk and recommendation if considered high/medium risk.</i>
	Medium risk			
	Low risk			
Publication Type	Public-use only	The package contains public-use data only. Codebooks and analysis code files will allow for verification of evaluation results.		<i>Please state for this data package: will there be public-use data only, restricted-use data only, or both and provide justification as this relates to enabling verification of evaluation results and/or broad usability of the data.</i>
	Restricted-use only			
	Both			

Formatted: Font: Bold

Formatted: Font: Bold

Data Component #3 (Panel Interviews) Preparation Overview

	Response		Discussion/Explanation
Data + Code Completeness	Complete		<i>To be considered Complete: The available data must allow new users to replicate evaluator analysis to the extent allowable by providing the full data set + analysis code. The constructed variables may also be included in a dataset, but if the dataset+code produces those variables, it is not necessary.</i>
	Incomplete	The data (as .csv) and R code (as .rmd) are complete and are contained in this package. Also created to support reproduction are a code book.	<i>To be considered Incomplete: The available data only provides a sub-section of data as produced by the survey and/or the constructed variables only. Incomplete data files are limited in terms of full verification of analysis and/or broad usability of data and must be justified.</i>

Formatted: Font: Bold

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

Data Round(s):	Baseline only	Combination of rounds. Contained in this package both the initial and final rounds of data collection, which occurred in August/September 2017 and August/September 2018, respectively.		<i>MCC is willing to trade-off broad use of individual rounds for more consistent de-identification protocols across rounds of data. Therefore, unless there is specific demand for the baseline/interim only data, or contractual requirements, MCC prefers contractors to prepare all data rounds in one package.</i>
	Interim only			<i>If one stage only – please (i) confirm demand and/or contractual justification and (ii) discuss how preparation and release of this data as presented to the DRB may affect future data round releases.</i>
	Endline only			<i>If combination, please discuss if this file replaces any previously published datasets.</i>
	Combination of rounds			<i>If combination, please discuss if this file replaces any previously published datasets.</i>
Informed Consent and IRB	High restriction	Given the nature of questions in the panel survey and the study overall, low restriction was required. As such and following expectations laid out in the consent script, only direct personal identifiers were removed from the dataset. The informed consent script states “TASCHA and MCC will remove any identifying information from the transcripts and notes – such as names, dates, and specific locations.” In addition, given the general nature of the questions and large population geographies and sizes that surround the RSRCs, we determined a very low risk of de-identification from indirect identifiers.		<i>MCC assumes DIRECT identifiers are always removed from any public-use file. With this assumption: Please refer to the informed consent statement – does it require: High restriction: access to data that includes indirect identifiers is limited to the contractor only; Medium restriction: access to data that includes indirect identifiers is limited to the contractor and qualified researchers, including MCC; Low restriction: data with indirect identifiers may be made public.</i>
	Medium restriction			<i>Please discuss how the promises of confidentiality in the informed consent informed de-identification efforts. Please include any additional guidance provided by the IRB as applicable.</i>
	Low restriction			<i>Please discuss how the promises of confidentiality in the informed consent informed de-identification efforts. Please include any additional guidance provided by the IRB as applicable.</i>
Geographic Identifiers	Highest (Region)	Oshakati (pop. 182K) Omaheke (pop. 71K) Ohangwena (pop. 245K)	The regions are so large and populated that individual identification through this variable is considered	<i>Please provide justification on the identification/de-identification/complete removal of specific geographic regions. De-identifying at a higher geographic level may support privacy protection, but it may also reduce data usability. Please provide justification for recommendation.</i>

Formatted: Font: Bold

Formatted: Font: Bold

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

			nearly impossible.	
	--(i.e. District)	NA – Not collected		
	--(i.e. State)	NA – Not collected		
	--(i.e. Village)	NA – Not collected		
	Lowest --(i.e. Census Blocks)	NA – Not collected		
Knowledge of Treatment	High risk	Knowledge of the survey taking place at the regional library is assumed to provide little to no re-identification risks. Given the number of patrons that use each library every year and the overall non-personal nature of survey questions, we deem re-identification a very low risk.		<i>In some cases, general knowledge of treatment areas and/or inclusion of a treatment variable can significantly increase re-identification risk depending on the population affected. Please provide assessment of this re-identification risk and recommendation if considered high/medium risk.</i>
	Medium risk			
	Low risk			
Publication Type	Public-use only	The panel interview data will be public-use only. Data dictionaries and analysis code files will allow for verification of evaluation results by enabling the user to repeat analysis queries.		<i>Please state for this data package: will there be public-use data only, restricted-use data only, or both and provide justification as this relates to enabling verification of evaluation results and/or broad usability of the data.</i>
	Restricted-use only			
	Both			

Formatted: Font: Bold

Formatted: Font: Bold

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

Section 3:

Data Component #1 (Patron Survey) Preparation Detail

Specific Issues		Risk Analysis		Risk Mitigation	
		Instructions	Response	Instructions	Response
1.	Who has significant financial, legal, cultural, or other incentives to re-identify survey respondents?	List all potential threats[1]	None. Given the nature of the patron survey, we do not foresee any meaningful incentives to for de-identification. These data cannot be connected to other studies that we are aware of, nor do we see individual participation in the study or use of the RSRC as a tool for embarrassment of the US government or MCC.		
2.	What is the potential value to these intruders?	List all uses (for example: capture delinquent tax payments, or stigmatize the respondent)	None. The data are not of a personal nature and there are no cultural stigmas associated with using the RSRC, nor are the data collected of such a personal		

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

			nature in the context of Namibia that they could be used to stigmatize the respondent.		
3.	What is the expected cost to these intruders to re-identify the data?	<i>Describe degree of difficulty for re-identification</i>	The cost would be very high as it would require significant time to go through the dataset, as well as do additional primary data collection to create linkages.		
4.	Assess availability of 'linkage' data that can be used to re-identify respondents. This includes other datasets or archives with information that can be used to re-identify individuals in the dataset.	<i>List all potential existing data</i>	It is nearly impossible for existing data to provide linkages. Although census data might have similar demographic information, raw census data are not available to the public, nor do they contain personal identifiers. In addition, given the population size of each region, we do not expect demographic	<i>Describe how to mitigate link to existing data that enables re-identification</i>	The risk of re-identification is so low that established protocols of removing direct identifiers is sufficient.

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

			responses to this survey unique enough for re-identification.		
5.	Identity Disclosures: What are the DIRECT identifiers in the raw data?	<i>List the DIRECT identifiers (names, addresses, geographic information, government-issued ID numbers, etc.)</i>	<ul style="list-style-type: none"> · Name · Phone number · Email · Additional contact information 	<i>List all DIRECT identifiers removed from the dataset.</i>	<ul style="list-style-type: none"> · Name · Phone number · Email · Additional contact information
6.	Attribute Disclosures: For GIS/GPS data, this distance data can be a direct identifier that is VERY useful analytically. Therefore, please describe how GIS/GPS data VALUE/USABILITY can be retained.	<i>List all GPS and/or GIS data.</i>	None. GIS/GPS data not collected.	<i>Describe process for de-identification. For example: introduce random errors into geographic data (GPS, GIS, etc.).</i> <i>Displace urban points 0-2 km, rural points 0-5 km, and additional 1% of rural points 0-10 km[2].</i>	NA
7.	Attribute Disclosures: What variables have OUTLIERS that create INDIRECT identifiers are in the raw data?	<i>List the identifying items/variables</i>	Although there are variables with outliers that have the potential for indirect identification, we see the possibility or re-identification as very low. Variable and response for each listed below.	<i>Describe top/bottom coding: set upper & lower bounds to remove outliers for continuous. Specify: are values set to the median, or other?</i> <i>For large categories/datasets, the OMB suggests top coding at least the highest .5%; for smaller categories/datasets, top</i>	NA

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

			<p>- <i>Frequency of using the RSRC and/or its services</i> Although frequency of use outliers, particularly on the high side, pose a potential for re-identification, we see that risk as very low. The survey was conducted 2 years ago, limiting recall. In addition, only staff would be reasonably expected to know who the extreme users were, and there are enough of those users within the overall RSRC-user population that re-identification would be very difficult.</p> <p>- <i>Household income/household amenity ownership</i> Household income was provided categorically, with the upper limit NA\$50K and above. This lack of detail significantly limits potential re-identification. Household amenity ownership included a few variables with low</p>	<p><i>code the highest 3-5%. The same principles apply to bottom coding.[3]</i></p>	
				<p><i>Describe any variables that require collapse and describe construction of new variable</i></p>	None
				<p><i>Describe any global re-coding to group observations into categories (e.g., age 0-5, 5-10, 65+, etc.). Ensure that the categories are neither too broad nor too narrow.</i></p>	None

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

			<p>response rates, including motorbike, landline phone, and internet connection. However, given the size of the populations surrounding the RSRCs, we do not see reidentification from these variables as a significant risk.</p> <p>- Time necessary to travel to/from the RSRC.</p> <p>The RSRCs were, by design, intended to serve geographically diverse populations. As such, there is an expectation that some (although the fewest) individuals would live far away. With the large size of the surrounding population, and given that distance is a rounded estimate, we see it unlikely that re-identification would be possible.</p>		
8.	<p>Attribute Disclosures:</p> <p>What variable combinations produce UNIQUE observations that create INDIRECT</p>	<p>List the identifying items/variables:</p>	<p>Although there are potential combinations of variables that could be used for re-identification (listed below), we</p>	<p>For each identified rare data, describe the local suppression techniques employed to remove unique and rare data.</p>	<p>NA.</p>

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

	IDENTIFIERS (for example: individuals with high incomes, ages, or unique combinations, such as 17-year old widowers or contextually unusual racial/ethnic backgrounds)		<p>see the risk as very low. This is because 1) the surveys took place in 2017 and 2018 so recall and personal changes of users limit knowledge of individual characteristics, 2) some variables are categorical so lack the specificity for re-identification, and 3) the size of the population surrounding the RSRCs is so large that many individuals could potentially have similar characteristics within the variables noted.</p> <ul style="list-style-type: none"> · Age · Occupation status · Primary occupation status (e.g. self employed) · School level and type (e.g. NAMCOL, NUST) · Number of people in household · Frequency of using the RSRC and/or its services · Household income / household 	Specify: are values set to missing, the median, or other?	
--	--	--	--	---	--

Commented [GAA(4)]: Spell out.

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

			amenity ownership · Time necessary to travel to / from the RSRC		
--	--	--	---	--	--

Data Component #2 (Panel Survey) Preparation Details

Specific Issues	Risk Analysis		Risk Mitigation	
	Instructions	Response	Instructions	Response

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

1.	Who has significant financial, legal, cultural, or other incentives to re-identify survey respondents?	<i>List all potential threats[1]</i>	Given the nature of the panel survey, we do not foresee any meaningful incentives to de-identification. These data cannot be connected to other studies that we are aware of, nor do we see individual participation in the study or use of the RSRC as a tool for embarrassment of the US government or MCC.		
2.	What is the potential value to these intruders?	<i>List all uses (for example: capture delinquent tax payments, or stigmatize the respondent)</i>	None. There are no cultural stigmas associated with using the RSRC, nor are the data collected of such a personal nature in the context of Namibia that they could be used to stigmatize the respondent.		

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

3.	What is the expected cost to these intruders to re-identify the data?	<i>Describe degree of difficulty for re-identification</i>	The cost would be very high as it would require significant time to go through the dataset, as well as do additional primary data collection to create linkages.		
4.	Assess availability of 'linkage' data that can be used to re-identify respondents. This includes other datasets or archives with information that can be used to re-identify individuals in the dataset.	<i>List all potential existing data</i>	It is nearly impossible to link data for re-identification. Although census data might have similar demographic information, raw census data are not available to the public, nor do they contain personal identifiers. In addition, given the population size of each region, we do not expect demographic responses to this survey unique enough for re-identification.	<i>Describe how to mitigate link to existing data that enables re-identification</i>	The risk of re-identification is so low that established protocols of removing direct identifiers is sufficient.

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

5.	Identity Disclosures: What are the DIRECT identifiers in the raw data?	<i>List the DIRECT identifiers (names, addresses, geographic information, government-issued ID numbers, etc.)</i>	<ul style="list-style-type: none"> · Name · Email · Phone number · Additional contact information 	<i>List all DIRECT identifiers removed from the dataset.</i>	<ul style="list-style-type: none"> · Name · Email · Phone number · Additional contact information
6.	Attribute Disclosures: For GIS/GPS data, this distance data can be a direct identifier that is VERY useful analytically. Therefore, please describe how GIS/GPS data VALUE/USABILITY can be retained.	<i>List all GPS and/or GIS data.</i>	NA. GIS/GPS data not collected.	<i>Describe process for de-identification. For example: introduce random errors into geographic data (GPS, GIS, etc.).</i> <i>Displace urban points 0-2 km, rural points 0-5 km, and additional 1% of rural points 0-10 km[2].</i>	NA
7.	Attribute Disclosures: What variables have OUTLIERS that create INDIRECT identifiers are in the raw data?	<i>List the identifying items/variables</i>	<i>- Frequency of using the RSRC and/or its services</i> Given the size of the population, we expect only RSRC staff to know who the extent of users with outlying (primarily on the high side) use frequencies are. Given this and the fact that responses were self-reported over 1 year ago,	<i>Describe top/bottom coding: set upper & lower bounds to remove outliers for continuous. Specify: are values set to the median, or other?</i> <i>For large categories/datasets, the OMB suggests top coding at least the highest .5%; for smaller categories/datasets, top code the highest 3-5%. The same principles apply to bottom coding.[3]</i>	NA

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

			we do not see risk in re-identification.	<i>Describe any variables that require collapse and describe construction of new variable</i>	None
				<i>Describe any global re-coding to group observations into categories (e.g., age 0-5, 5-10, 65+, etc.). Ensure that the categories are neither too broad nor too narrow.</i>	None

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

8.	Attribute Disclosures: What variable combinations produce UNIQUE observations that create INDIRECT IDENTIFIERS (for example: individuals with high incomes, ages, or unique combinations, such as 17-year old widowers or contextually unusual racial/ethnic backgrounds)	<i>List the identifying items/variables:</i>	None	<i>For each identified rare data, describe the local suppression techniques employed to remove unique and rare data. Specify: are values set to missing, the median, or other?</i>	NA
----	---	--	------	--	----

Data Component #3 (Panel Interviews) Preparation Details

Specific Issues	Risk Analysis		Risk Mitigation	
	<i>Instructions</i>	<i>Response</i>	<i>Instructions</i>	<i>Response</i>

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

1.	Who has significant financial, legal, cultural, or other incentives to re-identify survey respondents?	<i>List all potential threats[1]</i>	Given the nature of the panel interviews, we do not foresee any meaningful incentives to de-identification. These data cannot be connected to other studies that we are aware of, nor do we see individual participation in the study or use of the RSRC as a tool for embarrassment of the US government or MCC.		
2.	What is the potential value to these intruders?	<i>List all uses (for example: capture delinquent tax payments, or stigmatize the respondent)</i>	None. There are no cultural stigmas associated with using the RSRC, nor are the data collected of such a personal nature in the context of Namibia that they could be used to stigmatize the respondent.		

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

3.	What is the expected cost to these intruders to re-identify the data?	<i>Describe degree of difficulty for re-identification</i>	The cost would be very high as it would require significant time to go through the dataset, as well as do additional primary data collection to create linkages.		
4.	Assess availability of 'linkage' data that can be used to re-identify respondents. This includes other datasets or archives with information that can be used to re-identify individuals in the dataset.	<i>List all potential existing data</i>	It is nearly impossible to link data for re-identification. Although census data might have similar demographic information, raw census data are not available to the public, nor do they contain personal identifiers. In addition, given the population size of each region, we do not expect demographic responses to this survey unique enough for re-identification.	<i>Describe how to mitigate link to existing data that enables re-identification</i>	The risk of re-identification is so low that established protocols of removing direct identifiers is sufficient.

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

5.	Identity Disclosures: What are the DIRECT identifiers in the raw data?	<i>List the DIRECT identifiers (names, addresses, geographic information, government-issued ID numbers, etc.)</i>	<ul style="list-style-type: none"> · Name · Email · Phone number · Additional contact information 	<i>List all DIRECT identifiers removed from the dataset.</i>	<ul style="list-style-type: none"> · Name · Email · Phone number · Additional contact information
6.	Attribute Disclosures: For GIS/GPS data, this distance data can be a direct identifier that is VERY useful analytically. Therefore, please describe how GIS/GPS data VALUE/USABILITY can be retained.	<i>List all GPS and/or GIS data.</i>	NA. GIS/GPS data not collected.	<i>Describe process for de-identification. For example: introduce random errors into geographic data (GPS, GIS, etc.).</i> <i>Displace urban points 0-2 km, rural points 0-5 km, and additional 1% of rural points 0-10 km[2].</i>	NA
7.	Attribute Disclosures: What variables have OUTLIERS that create INDIRECT identifiers are in the raw data?	<i>List the identifying items/variables</i>	None	<i>Describe top/bottom coding: set upper & lower bounds to remove outliers for continuous. Specify: are values set to the median, or other?</i> <i>For large categories/datasets, the OMB suggests top coding at least the highest .5%; for smaller categories/datasets, top code the highest 3-5%. The same principles apply to bottom coding.[3]</i>	NA

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

				<i>Describe any variables that require collapse and describe construction of new variable</i>	None
				<i>Describe any global re-coding to group observations into categories (e.g., age 0-5, 5-10, 65+, etc.). Ensure that the categories are neither too broad nor too narrow.</i>	None

NAMIBIA & PERFORMANCE EVALUATION OF THE
REGIONAL STUDY AND RESOURCE CENTERS (RSRC) ACTIVITY
TECHNOLOGY & SOCIAL CHANGE GROUP, INFORMATION SCHOOL, UNIVERSITY OF WASHINGTON
Prepared on: March 21, 2019

8.	Attribute Disclosures: What variable combinations produce UNIQUE observations that create INDIRECT IDENTIFIERS (for example: individuals with high incomes, ages, or unique combinations, such as 17-year old widowers or contextually unusual racial/ethnic backgrounds)	<i>List the identifying items/variables:</i>	None	<i>For each identified rare data, describe the local suppression techniques employed to remove unique and rare data. Specify: are values set to missing, the median, or other?</i>	NA
----	---	--	------	--	----

[1] As stated in NIST 2016, de-identification practitioners should assume that de-identified US government datasets will be subjected to sustained, world-wide re-identification attempts, and they should gauge their de-identification requirements accordingly. Although a specific dataset may not be seen as sensitive, de-identifying that dataset may be an important step in de-identifying another dataset that is sensitive. Alternatively, the adversary may merely wish to embarrass the US government agency or its partners. Thus, adversaries may have a strong incentive to re-identify datasets that are seemingly innocuous.

[2] ICF International, Demographic & Health Surveys

[3] Office of Management and Budget, Checklist on Disclosure Potential of Proposed Data Releases ([current link](#))