DataFirst

Secure Research Data Centre Procedures Manual

1. Introduction

This *Secure Research Data Centre Procedure Manual* aims to assist researchers interested in obtaining access to our Secure Research Data Centre (the Centre) at the University of Cape Town. We share data on our open data site at https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/central. These data have reduced detail to protect respondent confidentiality. However, this can limit their research utility. The goal of the Secure Centre is to maximise data resources by giving researchers secure access to highly disaggregated data. In the Centre researchers can access data not otherwise available to the academic community.

Access to data in the Centre is free and available to researchers at universities or other research institutions. Applications are assessed on (a) whether the applicant is a bona fide researcher and (b) whether their proposed research is possible using the data in the Centre. This manual outlines procedure for

(i) Discovering Centre data
(ii) Applying for access to the Centre
(iii) Working in the Centre
(iv) Bringing data into the Centre to link with our data
(v) Requesting and receiving final research output from your work in the Centre

Information and forms for the Centre are available at https://www.datafirst.uct.ac.za/services/secure-data-services

2. Discovering Centre data

The Centre holds sensitive or potentially disclosive data e.g. firm-level data or data with household-level GPS coordinates. Researchers can read about the these data on our open data site https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/central/about. Documents that will assist analysis of this data, such as questionnaires, are also available on the site.

3. Applying for access to the Centre

Researchers can apply for access to the data in the Centre by completing the *Accredited Researcher application* form available at https://www.datafirst.uct.ac.za/services/secure-data-services. Applicants provide their name and contact details, academic credentials and details of their home institution. They also provide details on their research project and the data required in the Centre, plus the envisioned project time-frame to determine their period of access. Applicants must send their completed, signed form to us at support@data1st.org the

Applications are assessed by the Centre administrator for completeness, and then sent for to the data depositor for approval. 5 weeks should be allowed for the accreditation process, as it may involve discussions with the Centre administrator and depositors. Researchers will be contacted by the Centre administrator through our support site once a decision has been made.

4. Attending our information session

Securing the data and accrediting researchers is only effective in preventing data confidentiality breaches if researchers understand the purpose of these measures. DataFirst therefore adheres to international best practice by requiring researchers to attend a 20-minute information session with the Centre administrator to complete their accreditation. Successful applicants will be contacted to agree on suitable dates for these sessions. The sessions are held at DataFirst's offices in the School of Economics Building at the University of Cape Town. During the session, the Centre administrator will go over policies and procedures in the Centre, take fingerprint readings for the Centre's biometric access system, and allocate researchers a project number and password to access their workspace on the secure server.

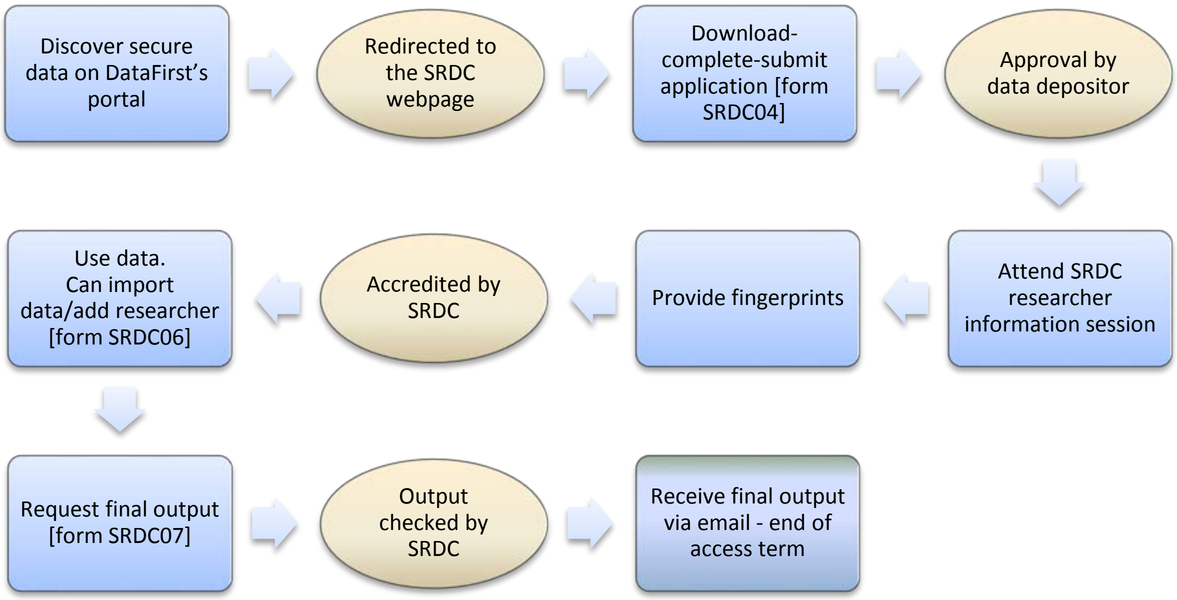
Discover secure data on DataFirst's portal
Redirected to the SRDC webpage
Download-complete-submit application [form SRDC04]
Approval by data depositor
Attend SRDC researcher information session
Provide fingerprints
Accredited by SRDC
Use data. Can import data/add researcher [form SRDC06]
Request final output [form SRDC07]
Output checked by SRDC
Receive final output via email - end of access term

Fig 1. Secure Research Data Centre Procedures

5. Analysing data in the Centre

After accreditation, researchers will be able to access the Centre with their fingerprint using the biometric reader. Researchers will be required to authenticate at login with their username and password provided during the information session. Researcher logins give access to a home directory and data analysis software. Home directories contain three folders: (i) a Working folder containing the data files and their approved application, and (ii) an Output folder which contains an *Output Request form* and where the researcher places the final output which they want to take out from the Centre.

6. Security in the Centre

Data may not be removed from the Centre and desktop workstations in the centre have had their BIOS secured, external ports disabled, and network properties shut off. Mobile devices and laptops may not be used in the Centre. Researchers can place these in the lockers provided for this purpose and the keys to these will be kept by the Centre administrator.

7. Software in the Centre

Software in the Centre includes
Adobe Reader
ArcGIS and QGIS
Microsoft Office Suite
R
SPSS
Stata
StatTransfer
Textpad
The use of other software can be accommodated.

8. Adding a research on a project

If a researcher joins a current project they will also need to apply for access. Their application should indicate that they are joining an existing project, and quote the project number allocated by the Centre administrator, as this will expedite approval.

9. Importing additional data

To add additional data to their workspace, researchers can

(i) Complete an online request form on our data site, if the data is public on our site, specifying the data should be added to their workspace on the secure server

(ii) Email their request and the data files to us at support@data1st.org if they need data from elsewhere

(iii) Email a request to us at support@data1st.org if they need additional restricted-access data

10. Extending access

Access to the Centre will expire 5 days after the last day of access specified in the researcher's Application. Their user account will also expire at this time. If researchers need more time on their project they can contact us at support@data1st.org before the project expiry date, and we will extend their access.

11. Requesting research output from the Centre

Allowed outputs from the Centre include tables, graphs, and syntax files. Log files may not be removed as the large amount of text in these files makes disclosure checking difficult. Researchers must place output they want from their research into the "output" folder in their home directory, complete the *Research Output Request form* in this folder, and email support@data1st.org to request their output. DataFirst must vet all files before they can be removed from the Centre. This is to ensure output from Centre is not disclosive (cannot be used to identify individuals). Checked output will be emailed to the researcher.

To avoid delays in receiving output, researcher should:

(i) Plan for the end of a project to build disclosure control time into final deadlines

(ii) Provide output files in as close to a final format as possible, easy to read and print, with clear headings and numbering, to avoid the reviewer having to spend excessive time trying to interpret them

(iii) Remove all references to individual identifiers in the submitted files

(iv) Keep content as brief as possible, as the reviewer will need to check *all* output, and extraneous items will slow down the procedure

NOTE: DataFirst reserves the right to charge for output checks of more than 30 pages, to ensure thorough output preparation on the part of researchers

12. Project completion and file deletion

Once researchers confirm their project is closed, the data and output will be deleted from the researcher's home directory. There is no time-limit to projects, and home directories will be left as is until project completion.

13. Citing the data in the Centre

Researchers agree to cite the data in the Centre in any published research. Correct citations for each dataset are available on their "landing page" on DataFirst's data site http://www.datafirst.uct.ac.za/catalogue3/index.php/catalog or researchers can contact support@data1st.org for advice on how to cite the data.

14. Copies of published research

As specified in the application, researchers must give DataFirst links to or digital copies of their research output from the Centre

15. Depositing variables for further analysis

DataFirst encourages researchers to deposit with us any new variables they create from their work in the Centre. These can benefit future researchers. Any non-disclosive derived data that has re-use value can be added to online versions of the data. If necessary, disclosive variables can be added to future versions of the restricted-access data. Researcher attribution for these variables will be included in the metadata.

SOURCES CITED

Altman, Micah. 2011. Managing Confidential Information in Research [Online]. Accessed 20120403 from http://www.slideshare.net/drmaltman/managing-confidential-information-in-research

National Research Foundation. 2008. NRF/SRS restricted-use data procedures guide. [Online]. Accessed 20121005 from http://www.nsf.gov/statistics/license/forms/pdf/srs_license_guide_august_2008.pdf

UK Data Archive Secure Research Data Centre. About us. [Online]. Accessed 20121005 from http://securedata.data-archive.ac.uk/about

United States Census Bureau. 2010. The Researcher Handbook: U.S. Bureau of the Census Center for Economic Studies Research Data Center's Handbook for Researchers. Washington: U.S. Census Bureau. [Online]. Accessed 20120502 from http://www.vrdc.cornell.edu/info747/Readings/Researcher_Handbook_20101101.pdf