



DataFirst Remote Access Data Enclave user Guide

Contents

1. Introduction	3
2. How our Enclave keeps data safe.....	3
3. How to find our restricted-access data	3
4. How to apply for Secure RADE access	3
4.1 Application form	3
4.2 RADE access and use requirements	4
4.3 RADE researcher information session	4
5. Logging on to the Data First VPN Group at UCT	4
6. Logging on to the Remote desktop environment.....	4
6. Working in your RADE remote desktop environment	7
6.1 Your Working Folder	7
6.2 Your Output folder:	8
6.3 Depositing new variables	8
6.3 Logging off	9
6.4 Working with project team members	9
6.5 Extending your RADE access	9
7. Software applications and other resources.....	9
8. Data citation and links.....	9
9. RADE researcher survey	9

1. Introduction

The Remote Access Data Enclave (RADE) was set up by DataFirst to give online but controlled access to restricted-use data held by DataFirst. The data is either more sensitive or more granular than the public access data we share on our [open data site](#). Our open data has reduced detail to protect the privacy of data subjects. However, this can limit the data's research utility. The aim of the RADE is to maximise data resources by giving bona fide researchers safe access to sensitive data not otherwise available to the academic community.

Access to data in the RADE is free and available to researchers affiliated to universities or other research institutions. Applications are assessed on whether the applicant is a bona fide researcher and whether their proposed research can be undertaken using the RADE data. This guide documents our RADE confidentiality and security measures as well as procedures for applying for RADE access, working in the RADE environment, and obtaining final output from your RADE-based research.

2. How our Enclave keeps data safe

Researchers are now able to work on our restricted access datasets from elsewhere, under certain conditions. The conditions comply with South African data privacy legislation (the [POPI Act](#)) and the [POPIA industry code of conduct for public universities](#). They are based on the [5 Safes](#) data protection framework of the [UK Data Service](#). Researchers are expected to comply with the rules related to this framework.



Safe Data

Direct identifiers are removed from the data



Safe Projects

Project proposals must be approved by datafirst



Safe People

Researchers at research institutions are eligible to apply and must sign a confidentiality agreement



Safe Settings

The remote access environment prevents unauthorised access and use and is monitored by digital supervision software



Safe Outputs

Research outputs are vetted before being sent to researchers

3. How to find our restricted-access data

You can find information on the data available in the RADE via the [Remote Access Data Enclave Collections page](#) on our data site. You can also read the metadata and download relevant documentation from the dataset landing-pages under this Collection.

4. How to apply for Secure RADE access

4.1 Application form

To apply for access, please send the completed RADE Application form to datafirst-support@uct.ac.za and attach a screenshot of your IP address to your application. Access is allocated at the Project level but each of your project team must also register. Please allow 2 months for your application to be processed. Approval may involve discussions between you and the RADE administrator and our data depositors.

4.2 RADE access and use requirements

To obtain access to our secure RADE environment you must agree that:

- You will attend a virtual RADE information session with DataFirst's RADE administrator
- The device you use for accessing the RADE will be one managed by your institution or a third party on behalf of your institution
- You have anti-virus and anti-malware software installed on your device
- You will not share your RADE logins with anyone
- You will ensure your user logins are not saved by your web browser
- The network interface used to access the RADE will be the only active service during your RADE session
- No screen capture tools will be running on the device during remote access
- Your screen will be locked when unattended
- You will not attempt to identify data subjects in the data

4.3 RADE researcher information session

Securing the data and research and approving 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 online information session with the RADE administrator to complete their access approval. Successful applicants will be contacted to agree on suitable dates for these sessions.

5. Logging on to the Data First VPN Group

If you are University of Cape Town staff you will be able to log in to the Data First Group on UCT's Virtual Private Network (VPN) with your institutional credentials.

As a non-UCT researcher you will need to apply for a Third Party Account to log in to the Data First Group on UCT's VPN. DataFirst will set up this account for you. Your Third Party Account number (T-Number) and VPN password will be provided to you at the information session.

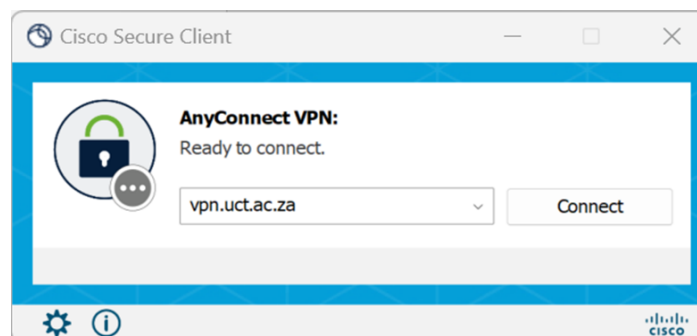
To log in to the Data First group, Third Party users must first download and install the Cisco (VPN) Agent from <https://vpn.uct.ac.za>. As an additional security measure, you will need to authenticate your login with the details sent to your cellphone. You must keep these logins safe and not share them with anyone, including members of your project team. If you forget your password, you can request a new password from datafirst-support@uct.ac.za

6. Logging on to the Remote desktop environment

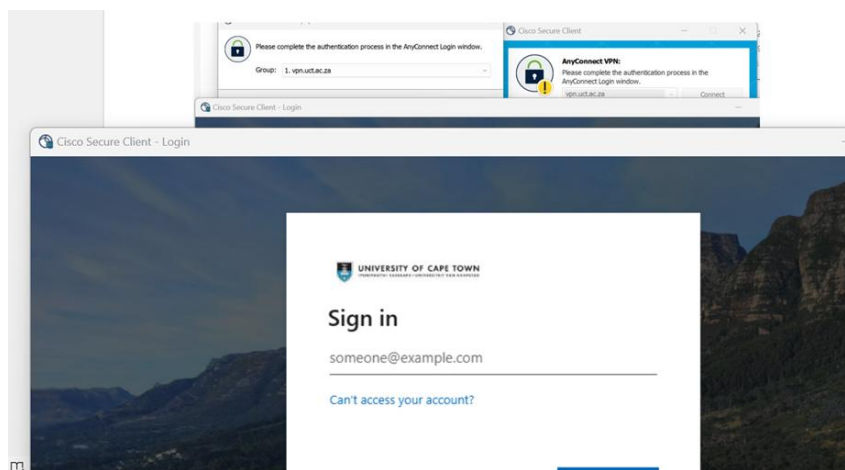
If your application is approved, your project will receive a RADE Project ID which will apply to all researchers working on your project. Each researcher will also be allocated a user login and password to log in to the remote access environment. Your RADE login details will also be provided to you at your information session.

Note: You cannot connect when on the UCT Eduroam network. You can use your home wi-fi connection or, when on campus, a cellphone hotspot connection to access the RADE (Settings-Connections-Mobile Hotspot) and connect to the Mobile Hotspot on your laptop/ desktop computer (Wi-Fi - select your phone connection). This will remove your Eduroam Wi-Fi connection.

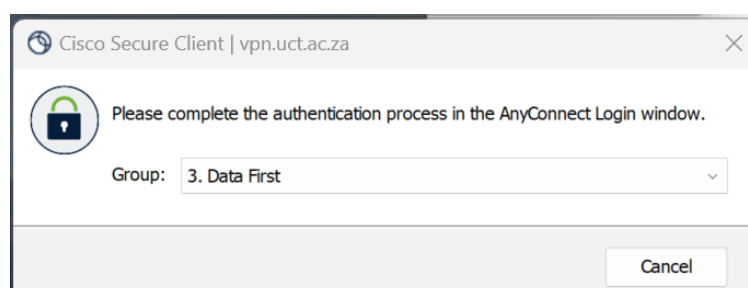
5.1 Log in to the Data First Group on UCT's Virtual Private Network by selecting the Cisco icon and selecting "Connect" on the VPN login page



You must move the Cisco login window aside to see the window with the options for Groups to connect to.

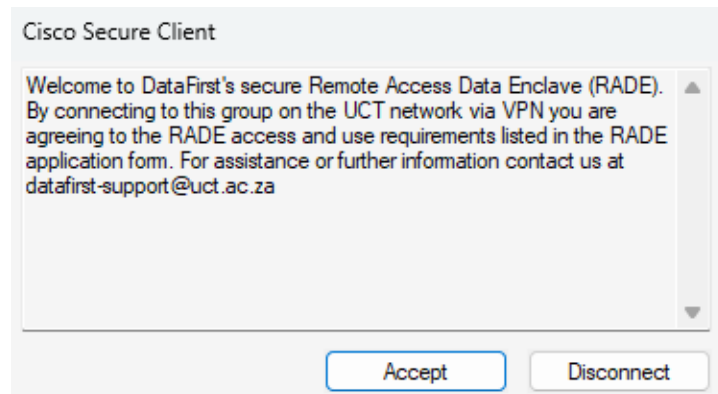


5.2 Select the Data First VPN Group



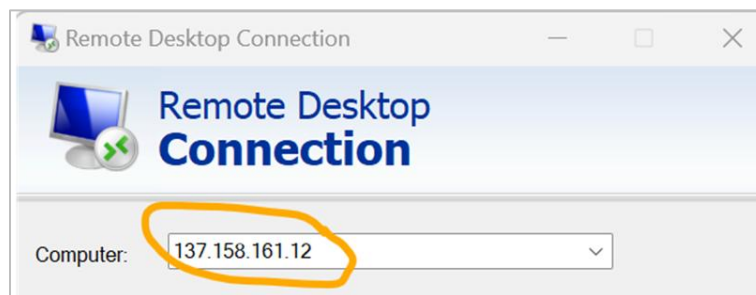
5.3 Authenticate your login either with the Authenticator app on your cellphone or the OTP that will be sent to you.

5.4 Accept the pop-up message with the RADE usage terms reminder and Select OK

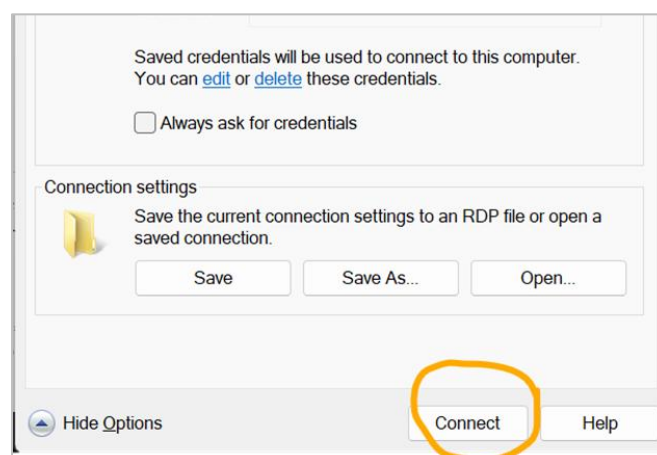


5.5 Select the Windows Remote Desktop Connection app (Settings–System–Remote Desktop)

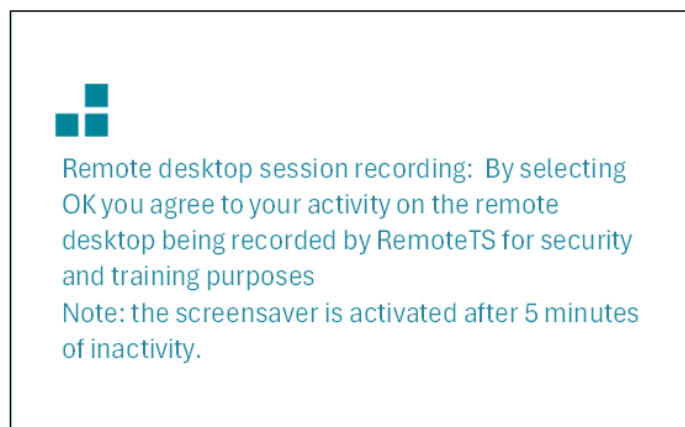
5.6 Select the RADE IP Address 137.158.161.12



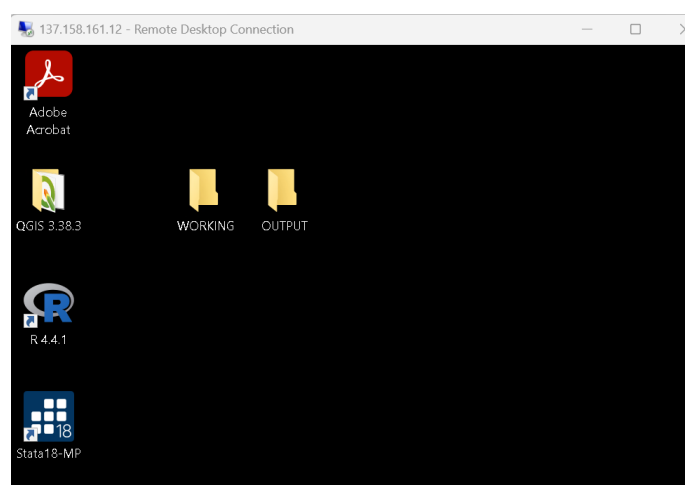
5.7 Select “Show Options” and Select “Connect”



5.8 Select OK to agree to the terms of the pop-up message that reminds you that your remote desktop session is being recorded for security purposes



You will now be logged in to your RADE remote desktop workspace account where you will see your Working and Output folders



6. Working in your RADE remote desktop environment

Your RADE account will give you access to your home directory and RADE software on the remote desktop. Your home directory contains two folders: a Working folder and an Output folder.

6.1 Your Working Folder

This folder is for you to save data files you are analysing, your draft publication and other items you are working on. This folder will contain a “Datasets” sub-folder where we will put the DataFirst dataset(s) you requested (data files and supporting documents). You must not write anything down from your RADE screen. You must not save files to your desktop as these will be deleted when you log off from the remote desktop session.

The Working folder also contains a “Syntax” sub-folder for your syntax files. You can send us any syntax files you would like us to place in your Working Folder. This is also where you must place the syntax files you create from your work in the RADE. The Syntax sub-folder is the only folder we will keep after your project closes off. We do not keep other files because space on the RADE is at a premium. This also means that if you overrun your space allocation you may be asked to delete some of your files/folders. But you can recreate your data using syntax files.

Your ‘Working’ folder includes a Data import request form. If you have data/ want to request data to be imported into your working folder other than our data you can complete this form and email it to support@data1st.org. You must ensure you have permission to use the data you want to import into the RADE environment. Please scan your data files for viruses or malware before providing them to us.

6.2 Your Output folder:

Your RADE home directory contains an Output folder. This is where you will place any output you wish to obtain from the RADE on completion of your research. Output means documents with statistical results prepared for publication or presentation, or syntax files. Data in any form may not be removed from the RADE environment. In this folder there is an Output request form for you to complete when you wish to remove output. The form provides DataFirst with information to assist us to undertake statistical disclosure control on your output to ensure it is safe to be released. Note the following regarding output:

- Your output must not be significantly different from that discussed in your original research proposal.
- You can only submit one output request at a time.
- You must submit a new output request for any output that was revised after changes requested by publishers/supervisors.
- If you would like to use the data you were granted access to for another project at close-off of your project you will need to apply for a new project

We do not release Microsoft Excel files, PowerPoint files or Stata log files because of confidentiality concerns. We also do not release intermediate results such as several tables/graphs with a data reference but no contextual text. This is because we need contextual information to undertake Statistical Disclosure Control on your RADE output and save the limited time of our output checker and to get your output released more rapidly. You can email us to inform us you have output in the folder. If the output is not disclosive it will be emailed to you.

6.3 Depositing new variables

DataFirst encourages you to deposit with us any new variables you create from your work on the data in the RADE. 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 restricted-access versions of the data. DataFirst will cite you as the creator of these variables in the citations for these datasets.

6.3 Logging off

It is important that you log off from the remote desktop properly to avoid future connection problems. Further, there are limited sessions on the RADE and therefore you should log off when finished so that you do not take up another researcher's remote desktop session. To log off properly, choose the logoff tab on the top right of your screen.

6.4 Working with project team members

Approved RADE researchers can work together on the same project. Researchers working on the same project will all be able to see their joint project folder. All project team members who wish to access the RADE must apply for access. Each team member will receive their own RADE login details. You must not share logins and only access the remote access with your own account details. New team members can be added at any stage by submitting an application, referencing the project number.

6.5 Extending your RADE access

Your RADE application specifies your project start and end date. If you wish to extend your access to the RADE please notify DataFirst a month before your project expiry date. Project access may be extended more than once.

7. Software applications and other resources

The remote desktop will give you access to the analytical and word processing software you need for your research. These include Adobe Reader, Microsoft Office, QGIS, QSR NVivo, R, SPSS, Stata, StatTransfer, and Textpad. If you need software application that is not installed please contact us to see if we can install it for you. On the remote desktop you will also find a folder called 'ADO Files' which will contain ADO files for Stata users. You cannot install Stata ADO files as downloads because there is no internet connectivity in the RADE. You must therefore run the ado files from this folder.

8. Data citation and links

You must cite the dataset and the RADE in any research you publish based on the RADE data. The correct citation format for each of our restricted-access datasets is available in the online metadata posted with each dataset in the [Remote Access Data Enclave Collection](#).

When you apply for RADE access you agree to send DataFirst links to any published research based on the restricted-access data in the RADE. Please follow up on this agreement by sending us the publications/links after publication.

9. RADE researcher survey

DataFirst will send you a survey form to complete after your project is completed. Please consider completing the survey as this presents an opportunity for you to tell us about your RADE experience. Your input will help us improve the RADE service for you and other researchers.

Appendix A. RADE flowchart

