

Training and Reference Guide

Welcome to the Data Enclave! This document will serve as an introduction to working with the data in the Enclave as well as a reference guide for continued use.

Requirements for Access

Please make sure your system meets the following requirements:

- Microsoft Windows 7 (Home, Professional, Ultimate edition) or higher or Mac OS
- Supported version of Internet Explorer (11 or higher), Safari, Firefox, or Chrome
- Citrix Receiver version 4.6 or higher: <https://www.citrix.com/go/receiver.html>
Note: You do not need to create a new “account” when installing a Citrix Profile
- Add the log in page enclave.norc.org to your trusted sites on internet browser

About the Enclave

The Enclave should be considered an extension of your research space. It is configured to provide you with all the tools you need to test your research questions, analyze data, prepare publications, and other output in a collaborative environment. The enclave provides resources to users on a virtual desktop. Some applications are installed locally on the desktop itself, and some are published to the desktop from separate, dedicated application servers.

Security

The enclave is a secure remote access environment for the analysis of confidential data. There are a number of security measures in place to protect the environment and to allow users to work with data in a safe manner. The most important of these is the fact that the enclave is a closed environment. This means that you cannot access the Internet or email from within the enclave and you cannot upload anything to or download anything from the environment to your local computer. You also cannot copy and paste text into and out of the Enclave.

Maintaining security also depends on users following best practices when accessing the environment. Please make certain to adhere to the following guidelines when working in the Enclave: do not share your password or username with other users and do not permit anyone who is not an authorized user to view the enclave. This includes but is not limited to permitting anyone from looking over your shoulder while you are logged in or attempting to share your screen or utilize screen capture software or devices.

If you have any questions about these requirements, please reach out to us at dataenclavemanager@norc.org

First Login

Log in Page: <https://enclave.norc.org>

Username: Emailed to you

Password: Emailed to you

Before you login, you should know that after three consecutive failed attempts to login will lock you out for 30 minutes. Your account will automatically unlock after 30 minutes, so there is no need to email the helpdesk unless you forgot your password.



Steps for first login:

1. Navigate to the Data Enclave login page. Enter your UserName and Data Enclave password in the first two fields.
2. On your first login, it may ask you to “Install” the Citrix Receiver, even if you already have. Please go ahead and finalize the installation by clicking “Install”.
3. When your Citrix Profile loads, click on Resnik Desktops which should launch your desktop.

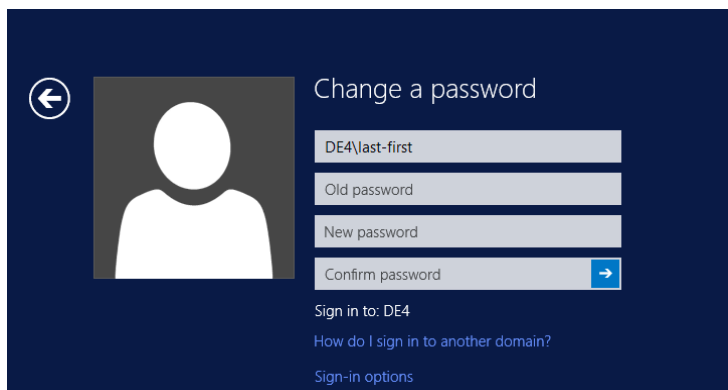


Change Your Password

You must change your password as soon as you log in for the first time. At the top of your Enclave window you should see a pull down menu with a downward pointing arrow:



There is a “ctrl+alt+del” icon. Click on this menu and select the “ctrl+alt+del” icon. A menu with the option to “change a password” will appear. Select this option. Enter the temporary password you used to log in as your old password.



Password Requirements:

8 characters minimum and at least:

- 1 uppercase letter
- 1 lowercase letter
- 1 number
- 1 symbol

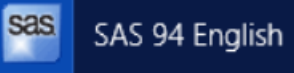



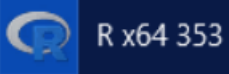

Your password cannot include any part of your name or a previous password that you have used in the Date Enclave.



Citrix Applications



From the start menu you can access your applications by clicking on the downward facing arrow in the bottom left corner of the desktop. Your account includes access to SAS, Stata, and R.

Please note that SAS limits queries to 50 million records or less. Also please note that you cannot open SAS files directly through the file explorer. You have to open the files directly from the SAS application. Similarly, you cannot open Stata files directly through file explorer. You have to open the files through Stata. Here is some basic information on each statistical application. Most common packages are already installed in the repository. We update applications on a regular schedule as needed. If you have other questions about the applications, please reach out to the Helpdesk at DataEnclaveManager@norc.org

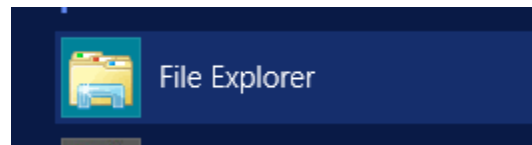
Application	Icon	Version	Description
SAS		SAS 94 English	SAS (Statistical Analysis System) is a software suite for advanced analytics, multivariate analyses, business intelligence, data management, and predictive analytics. You cannot open SAS files directly through the file explorer. You have to open the files directly from the SAS application. Please note that SAS limits query to 50 million records or less.
Stata	 	StataMP 14 64-bit; StataMP 15 64-bit	Stata is a general-purpose statistical software package. You cannot open Stata files directly through file explorer. You have to open the files directly from the Stata application. To install a common package from the repository, enter the 'net' command and click more to reveal all directories.
Stat/Transfer		Stat/Transfer 64-bit	Stat/Transfer is a utility program that converts data sets from one format to another (for example SPSS to Stata, or SAS to Excel).
R		R x64 353 (version 3.5.3)	R is a programming language for statistical computing and graphics supported by the R Foundation for Statistical Computing. To install a common package from the repository, select "Package" in the toolbar, and "Install package(s)".
RStudio		RStudio version 3.5.3	RStudio is an integrated development environment (IDE) for R. To install a common package from the repository, select "Packages" on the right-hand side, install and search for

			the package name in the pop-up “Install Packages” and search by the package name.
Jupyter Notebook		Python 3.6	Jupyter supports execution environments in several dozen languages; including Python, R, Rstudio, Scala, and MATLAB. To install a common package from the repository, enter the “ <u>import package name</u> ” command. If your desired package is not available, email the helpdesk to have it added.
Spyder		Python 3.6; Spyder 3	Spyder is an open-source cross-platform integrated development environment (IDE) for scientific programming in the Python language.

Data Areas

In the file explorer application, you should see two drives:

(1) DOCUMENTS: Your personal workspace. Only you and NORC administrators have access. When you request an import (see below), your import requests will appear in the import folder.



(2) RESNIK_RESEARCH: shared data drive with three subfolders:

- (a) “Datasets”: a shared space for you and Resnik staff/users who have producer accounts
- (b) “Home”: links to the “documents” drive.
- (c) “Researchers”: a shared common folder for all Resnik researchers and a folder for you and your research team (if you are working in a group)

See the below section for information on how to access the data for this shared task.

Accessing Mental Health Cluster:

1. You can access the AWS instance through PuTTY to run experiments.
 - a. Your PuTTY instance should have a saved session named ‘Cluster’ - double click to launch your SSH session.
 - b. If necessary, information on establishing an SSH connection can be found in HPCREADME.txt (on Desktop).
2. When on the Mental Health Cluster, you can install available software packages (e.g., using pip) and issue jobs using slurm. Information on this process is found in HPCREADME.txt. After submitting a job using slurm, you will see your remaining AWS credits.
3. The data and baseline code for the shared task is located under “~/teamdata” on the Mental Health Cluster.
4. Files can be moved to and from the Mental Health Cluster using the WinSCP application. Users should use the following setting, where the host name is updated to include your team number:
 - a. File protocol: SFTP
 - b. Host Name: mhdeteam##.de4.norc.org (replace ## with your team number)
 - c. Port number: 22
 - d. User Name: your Data Enclave username

- e. Under the advanced setting, SHH, Authentication, you need to give the path to your private key file. This file should be found on your desktop or in your Documents (H:).

Importing files:

1. E-mail request to dataenclavemanager@norc.org with the following information:
 - a. A brief description of each file for import
 - b. The size of each file for import
 - c. Whether or not the file contains sensitive and/or confidential information
 - d. Do NOT provide your file(s) for import on this email
2. **If the file contains sensitive and/or confidential information** be sure to password-encrypt your file(s).
3. A NORC Data Enclave Manager (DEM) will acknowledge your request, then send a separate email containing a link to upload files via Accellion kiteworks (first-time users will need to set up their account).
4. Upload your file(s) to Accellion kiteworks
5. NORC will save the file to your H drive and you will be notified when the import process is complete.
6. Allow for 1 business day.

Exporting files:

1. To launch DEER, open the Internet Explorer link on your NORC Data Enclave desktop.
2. Upload the file that you would like to export out of the Enclave. If you have more than one file, please zip them.
3. To help expedite the review process, please include details about the file:
 - A brief description of the output, e.g., simple regression, box-and-whiskers plot, etc.
 - Whether this is a new request or one that builds on a prior request
 - A list of the key variables, and
 - Whether or not the variables are continuous
4. Email dataenclavemanager@norc.org to notify NORC that you saved a file to DEER for review.
5. A data custodian at NORC will pass the file for disclosure review.
6. If the file is approved, you will receive an email with a link to the file so that you can download your file outside of the Enclave. If the file is rejected, a data custodian at NORC will email you and explain why the file was rejected. Once the appropriate edits are made, you can resubmit your file for review.
7. Please allow for 2-3 business days for your export request to be completed.

To help expedite the review process:

- Do not request observation level data, include record accounts, and avoid minimums and maximums
- Do not request any data where the number of observations behind any statistic is less than 10
- Limit volume to a few pages

Leaving the enclave:

It is best practice to save your outputs and sign out after each session, especially before the monthly maintenance window.

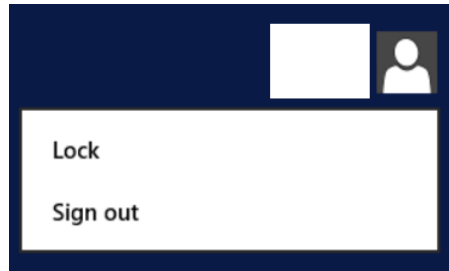
System Maintenance

System maintenance that impacts user access to Data Enclave desktops and applications will take place on the first Monday of each month from **2 am to 5 am (Central Time)**. During this time, all existing user sessions will be terminated. Please make sure that you have saved your files and logged out before the maintenance window.

Sign Out

- Go to the start menu. Click on your name at the upper right corner and you'll see a "Sign Out" option.

Disconnect



- If you need to run a program overnight, you can use the "Disconnect" option in the pull-down toolbar.
- *If your session is inactive for **72 hours**, it will be terminated.*
- Make sure to log back on so that it remains active. Make sure to save your outputs so they are not lost due to a terminated session.



For any questions or issues with the Data Enclave

Please contact the **Helpdesk** during business hours.

DataEnclaveManager@norc.org

9 am – 5 pm Central Time
Monday through Friday