MHDO Secure Data Access: NORC Data Enclave

Demo 2024

HSRI

Overview of the NORC DATA ENCLAVE

- What is the Data Enclave?

  - Secure, remote access to confidential data

- Initial launch sponsored and guided by the National Institute of Standards and Technology (NIST)

- Provides a robust analytic platform that enforces secure and compliant access to data

- Securely stored and provided access to MHDO data since 2012

Data Enclave Platform

Platform + Tools Approach:

- Fully managed infrastructure – users focus on analysis

- Draws upon existing skills and user communities for analytic tools and support

- Flexible/interoperable tools accommodate development and growth

Data Enclave
Security

- NIST 800-53 Rev5 compliance at the moderate control level
- Ensures ongoing maintenance of policies and procedures surrounding key areas of operations

TABLE 1: SECURITY CONTROL IDENTIFIERS AND FAMILY NAMES

ID	FAMILY	ID	FAMILY
AC	Access Control	MP	Media Protection
AT	Awareness and Training	PE	Physical and Environmental Protection
AU	Audit and Accountability	PL	Planning
CA	Security Assessment and Authorization	PS	Personnel Security
CM	Configuration Management	RA	Risk Assessment
CP	Contingency Planning	SA	System and Services Acquisition
IA	Identification and Authentication	SC	System and Communications Protection
IR	Incident Response	SI	System and Information Integrity
MA	Maintenance	PM	Program Management

Data Enclave User Management

- Microsoft Active Directory Services

- User account and permissions requests are processed in the Data Enclave using a ticketing system

- Account configuration details are captured in the ticket and implemented by IT staff

- **Views will be provisioned based on MHDO Data Request/Release Procedures**

Data Enclave Platform

Default Applications included in all Data Enclave accounts:

- R/RStudio, Stata, Python and Stat/Transfer
- Note: Python is supported by Spyder and Jupyter
- DBeaver
- Notepad++
- Firefox
- Microsoft Edge
- Acrobat Reader DC
- Microsoft Office 2013 (Word, Excel, PowerPoint)
- Windows Accessories (Snipping tool, Calculator, etc.)
- 7-Zip File Manager
- File Explorer (Stores flat files)

Requested Software Requiring Licensing Costs:

- ArcGIS
- GAMS
- GAUSS
- HLM
- Matlab
- SAS
- SPSS
- Tableau (User License Required)
- UltraEdit

Data Enclave User Information

Specifications	Direct Data Access in Data Enclave Analyst (Standard Subscription Fee)	Direct Data Access in Data Enclave Custom Project (DE Estimated Fee)
Data Access & Formats	Access database tables/views in Vertica HP (read only)	On request, requiring Feasibility Assessment and Estimating by NORC
Storage for data files, reports, spreadsheets, documents	50 GB personal H Drive: 200 GB shared storage available for the research team	Custom Database Storage + Custom File Storage as agreed with NORC
Database Software	Vertica HP database with Dbeaver SQL Client Tool	Vertica HP database with Dbeaver SQL Client Tool + Other Data Management Tools if supported by NORC Data Enclave
Data Analysis Software	STATA, R Studio, Python Access to Tableau Desktop with provision of external license	SAS, SPSS, ArcGIS+ Other Data Analysis Software, if supported by DE For Additional Software, customer may need to provide license; software must be Citrix compatible
Presentation Software	Microsoft Office Productivity Suite - Excel, PPT, Word, etc.	Microsoft Office Productivity Suite - Excel, PPT, Word, etc.
Download Files/Reports	5 SDC-reviewed export requests per month (Aggregated Data files, Reports)	Custom Number of Downloads and Size, as agreed with NORC
Upload 3rd Party Data	50 GB for Ancillary Data Files	Custom Number of Uploads, as agreed with NORC
Support Tier Description	2-day response service level agreement (SLA) for technical issues, defined as: account management, application launch failure, and database connectivity errors.	1 day response service level agreement (SLA) for technical issues, defined as: account management, application launch failure, and database connectivity errors.
Training	Self-guided onboarding to Data Enclave/MHDO Data via FAQs and short web-based video. 1-hour initial environment and data access training	Self-guided onboarding to Data Enclave/MHDO Data via FAQs and short web-based video. 1 hour initial environment and data access training
Data Objects Permissions	Read only access to data views for all available years in Vertica (low memory and temp table limits)	Tables/Views can be migrated to PROD (persistent) by NORC IT only
Database New Objects development	Creation of Temporary tables in Vertica is allowed	Creation of Temporary tables in Vertica is allowed
Request Process	Request and Approval by MHDO/and NORC Data Enclave Database Managers	Project Request and Approval to MHDO/scope & feasibility assessments by NORC
NORC Fee Structure	Hosting Fee Per User/Per Year = $5,356	Hosting Fee Per User/Per Year = $5,356 1 TB Vertica Storage ~ $10K / 1 TB 1 TB File Storage ~ $2,000 / 1 TB

USM Questions

1. **How is it structured? Would someone have their own working folder where once could create working datasets from the APCD to use.**

   Yes, users will have both their own folder and there can be a shared group folder with other USM staff.

2. **Is there SAS ,STATA. EXCEL, WORD on the environment to create working files and do analysis?**

   Yes, these tools are all available. SAS does have an additional monthly fee of $164/month.

3. **Would others be able to have access to say a share folder where the project was and the data set etc.?**

   Yes

4. **If one created tables to be used in a paper for example would they be able to be downloaded and output to an external USM folder or somewhere to add into a working paper within a USM computer?**
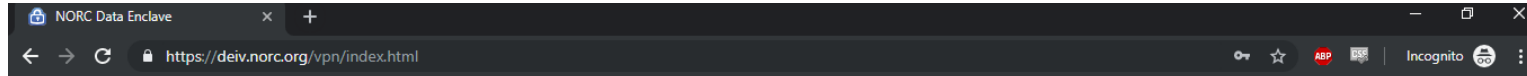
   As long as the DUA with MHDO allows for this.

5. **Is there a limit on data size ?**

   There are limits per account, but additional space can be purchased.

6. **What hours is the enclave available to use?**

   24/7 with schedule maintenance windows.

Data Enclave Demo