

# Diverse Interdisciplinary Research Teams and Secure Access To HIPAA Protected Data

Dr. D. Gardner, Dr. M-A. Demuynck, (Mathematics and Computer  
Science), Dr. M. Tietze (Nursing)

# Background & Context

- Data
  - Mental Health Data Spanning 5 Years
  - Sourced from DFWHC Foundation Warehouse
  - Diagnosis (Primary or Secondary) of Schizophrenia (ICD-10 Code of 295)
  - Administrative Hospitalization Information, Demographics, Financial Data, Predictive Information Such as Risk of Mortality Scores and Severity of Illness scores

# Background & Context

- Research Team
  - Interdisciplinary/Interprofessional with an Informatics Focus
  - Nursing, Health Science, Library Science, Data Science, Computer Science, Business
- Environment
  - Small to Medium Sized University
  - Faculty & Staff on Research Team Spread Across 3 Campuses

# Obstacles

- Data Set of Over 870,000 Records and 120 Fields
- Flat File Stored on Windows Network Share
  - Excel Spreadsheet
  - File Server Housed Shares for Multiple Individuals & Groups
- Slow to Load, Slow To Manipulate, Just Slow
- Needed to be Housed in a HIPAA Compliant Database
  - No Funding for Hardware or Software
  - No Funding for Setup & Administration Support
  - Encoding and Data Characteristics Made it Difficult to Import Into a Database

# Solutions

- Data Server
  - Perl & Python Scripts Along With Open Source Tools Used for Data Conversion & Migration
  - Debian Linux VM on Isolated HIPAA VM Farm
  - MariaDB Service Configured To Plugins for Encrypted Data
  - Secondary Encrypted Database as Backup of Original Data Set
  - Primary Database Allows For Table Creation by “Designers” Only
  - Multiple Frontends Allowed (MySQL Workbench, HeidiSQL, SAS)

# Solutions

- Access Control
  - Network Level Firewalls To Isolate VM 'Farm'
  - Research Team and Research Lab Machines Assigned Static IPs for Machine Based Firewall Control
  - Only Port 3306 Open For Researchers & Port 22 For System Administrator
  - TLS Encryption Required Across Port 3306
  - Database Accounts Tied To Static IP Except in Lab
  - Database Roles Used to Assign Privileges

# Improvements

- Database Design & Management
  - Break Up Single Table as Needed and Conduct Further Data Cleanup
  - Automatic Backups of Primary Database to Alternate Server
  - Creation of Key Server to Support Rolling Ciphers
- Access Control
  - Client Certificates for Researcher and Lab Machines
  - Researcher Machine Setup Script

# Improvements

- Automation
  - Automatic Key & Certificate Creation Scripts
  - Database & Plugin Scripts To Setup Encryption on Future Databases
  - Database Role and Account Generation Scripts Relying on User Data File
  - Base Server Platform Scripts To Automate Setup & Package Install
  - Master Script
- Training/Documentation
  - Expansion of Existing Access Training & Documentation