

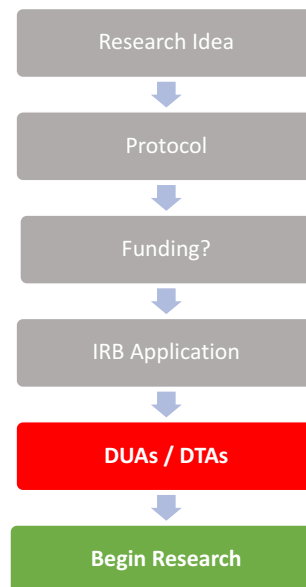
The Basics of VHA Data

Accessing, Requesting, and Analyzing

Laura A. Graham, PhD, MPH
Health Services Research Fellow, Ci2i, VA Palo Alto
Postdoctoral Fellow, SPIRE, Stanford University

Prior Steps

- Before requesting data access...
 - **Write a clear protocol**
 - **Know your study team**
 - **Get VA Access (Employed, WOC'ed)**
 - **Get IRB Approval!**



Seeking Information on VA Data?

- VHA Data Portal
 - <http://vaww.vhadatportal.med.va.gov/>
 - Great resource for VHA data
- VHA Data Architecture Repository
 - <https://vaausdarmul81.aac.dva.va.gov/>
 - Detailed data dictionary for CDW tables
 - Search string can be useful for finding data buried in CDW tables

Seeking Information on VA Data?

- VA Informatics and Computing Infrastructure (**VINCI**)
 - <https://vaww.vinci.med.va.gov/vincicentral/>
- VA Information Resource Center (**VIREC**)
 - <http://vaww.virec.research.va.gov/>
- VA Health Economics Research Center (**HERC**)
 - <http://vaww.herc.research.va.gov/>
- Office of Performance Measurement (**OPM**)
 - <https://vaww.car.rtp.med.va.gov/>
- Planning System Support Group (**PSSG**)
 - <http://vaww.pssg.med.va.gov/>
- Data.gov / Corporate Databases Monograph
 - www.data.gov

Readily Available Data Sources

- VHA Support Service Center
 - <https://vssc.med.va.gov/>
 - Facility-level information
 - Business Operations
 - Clinical Care
 - Quality & Performance

- National Center for Veterans Analysis and Statistics
 - <https://www.va.gov/vetdata/>



DART

Data Access Request Tracker

Requesting VHA Data Access

- Most of your data access requests will be covered by the **Data Access Request Tracker (DART)**
 - <https://dart.vha.med.va.gov/>
- Other data sources exist and their request processes can be found in the **VHA Data Portal**
 - <http://vaww.vhadatportal.med.va.gov/>

Data Access Request Tracker (DART)

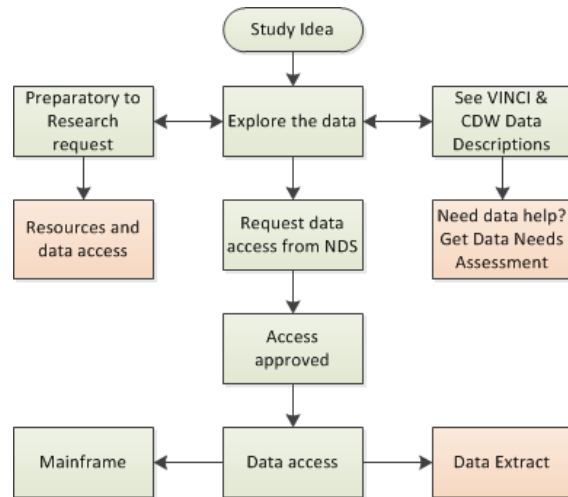
- Online workflow application
- Data requests and Prep to Research Requests

Showing 1 to 5 of 5

Tracking #	Name	Type	Contact	Submitted	Status	Actions
2012-09-0044-D-A02	Cardiac Risk and Silent Effect on Adverse Perioperative Events	Research Data Access	Hawn, Mary T	08/28/2014	NDS APPROVED	Actions
2014-02-027-D	CAPRI and Vista/Web	Research Data Access			NOT SUBMITTED	Actions
2014-10-096-D	Mitral Valve Operations: Trends, Effectiveness, and Antithrombot	Research Data Access	Hawn, Mary T	11/05/2014	NDS APPROVED	Actions
2014-12-124-D-A04	Improving Surgical Quality	Research Data	Morris, Melanie	01/27/2017	NDS APPROVED	Actions

Data Access Request Tracker (DART)

- DART is your starting place for VINCI data access among other sources
- Acquires National Data Service (NDS) approval
- Once approved an assigned VHA data manager will provide a VINCI Workspace and access to data
- Will help you navigate VINCI



VINCI helps with orange boxes

Data Access Request Tracker (DART)

- <https://dart.vha.med.va.gov/>
- Complete and upload the required documents as listed in DART for the data source(s) or data tool(s) selected.
- Some of the following documents will be required:
 - **Research Project Documents and Approvals**
 - Research study protocol
 - Research and Development Committee Approval Letter(s)
 - Institutional Review Board (IRB) Approval Letter(s)
 - IRB Approved HIPAA Authorization or [Waiver of HIPAA Authorization](#)
 - IRB Approved Sample Informed Consent or Waiver of Informed Consent

Data Access Request Tracker (DART)

- Additional documents that may be required:

- **Data Request Forms**

- [Research Request Memo](#) – Replaces the 9957 for **ALL** research requests submitted through DART after March 25, 2015
 - [Real SSN Access Request](#) – Required for access to real SSNs
 - [CDW Domain Check List](#) – CDW data only
 - [Vital Status Rules of Behavior](#) – Vital Status File only
 - [Special User Access Request Form for Researchers](#) – CAPRI and VistAWeb only
 - [Surgery Data Research Proposal Template](#) – Surgical Quality Data Use Group (SQDUG) data only
 - [National Surgery Office \(NSO\) Data Use Agreement and Security Information](#) – SQDUG data only

Data Access Request Tracker (DART)

The screenshot displays the DART web application interface. At the top, there is a blue header with the VAND logo and the text 'Data Access Request Tracker (DART)'. Below the header, a navigation bar shows 'HOME > DART Dashboard > Requests'. The main content area is titled '2014-12-124-D-A04 | Improving Surgical Quality: Risks and Impact of Readmission'. A blue bar indicates 'AMENDMENT NARRATIVE: Add Personnel'. Below this, the 'Activity Information' section is visible, containing a 'REQUEST INFORMATION' form with the following fields:

Short Name *	Improving Surgical Quality: Risks and Impact of Readmission
IRB Number *	CIRB: 14-07
IRB Expiration *	11/03/2017
Start Date *	11/04/2014
End Date *	04/30/2018

At the bottom of the form, there are buttons for 'Previous', 'Validate', 'Save Draft', and 'Next -->'. A small asterisk indicates that the fields are required.

Data Access Request Tracker (DART)

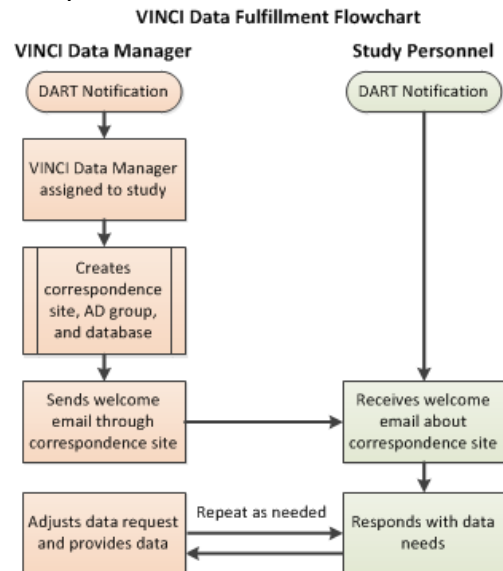
- Once all required documents are **uploaded** and the request has been **submitted**, it will undergo a privacy review.
- Additional reviews are required for the following:
 - Requests for data with real SSNs will undergo review by the Office of Research & Development.
 - Requests will undergo a security review using the information provided in the Research Request Memo.

Data Access Request Tracker (DART)

- When do I need to request **Real SSNs**?
 - You're planning to use chart abstraction
 - You're linking to another existing data source by SSN
 - You're planning to use the TIU (Text Integration Utilities)
- Everything else is linked by either **Sta3n / PatientSID** or **SCRSSN** and you will receive a crosswalk file once your request is approved

Data Access Request Tracker (DART)

- **After data access approval** a VINCI Data Manager can help with the following:
 - Cohort selection
 - Data Extraction
 - Formatting, Indexing, and Security



Other Data Access Examples

- Office of Reporting Analytics Performance Improvement and Deployment (**RAPID**)
 - DUA Forms
 - Project Information Sheet
 - Personal Agreement Statement
 - Data Access List
 - Study Protocol
 - IRB Approval
 - Data Request Specifications
- Final documents are submitted via email to CARDUARrequest@va.gov

VINCI

The VA Informatomics and Computing Infrastructure

VINCI Introduction

- Web-based platform for accessing and analyzing data with a variety of tools
- Provides consultation services for IRB Research for the entire life-cycle of a research project
- <https://vaww.vinci.med.va.gov/vincicentral/>

The screenshot shows the VINCI Central website interface. At the top, there is a navigation bar with links for Home, About VINCI, and Skip to Content. The main header area includes the VINCI Central logo, a welcome message "Welcome to the VA Informatomics and Computing Infrastructure", a keyword search box with a Search button, and contact information for Laura Graham, including her email (VINCI@VA.gov) and a Concierge Desk phone number (1-801-872-3324) with a link to Request Help Online.

Below the header, there is a sidebar menu on the left with the following items:

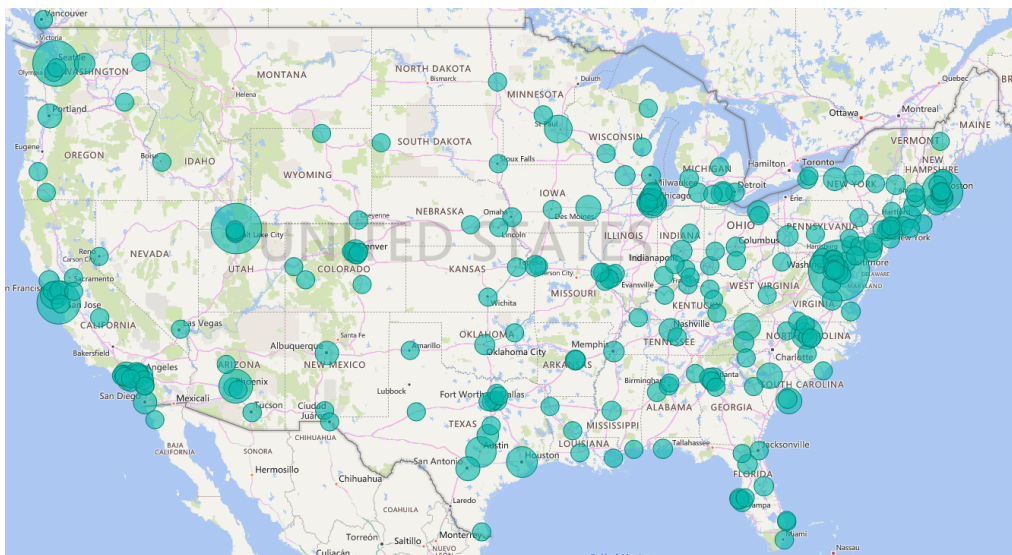
- Launch Workspace
- New To VINCI
- My VINCI Dashboard
- VINCI University
- VINCI Tube
- VINCIPedia/FAQ
- User Guides
- VINCI Services
- Computing Cloud

The main content area features a "New to VINCI" banner with the text "Learn more about VINCI, VINCI Studies and the virtual environment." and a "My VINCI Dashboard" section with the text "All your VINCI resources in one place!". At the bottom, there are four quick links: VINCI University, VINCI Tube, VINCIPedia, and User Guides.

VINCI Introduction

- <https://vaww.vinci.med.va.gov/vincicentral/>
- 105 high-performance servers and 1.5 petabytes of data storage
- 450+ new projects – 17% increase YOY
- Provided data to 840+ projects
- VINCI Data Analysts touch about 50 projects/week
- Structure
 - Windows 2012R2 Workspace
 - SQL Servers (Currently 3)
 - **RB01:** vhacdwwbs01.vha.med.va.gov
 - **RB02:** vhacdwwbs02.vha.med.va.gov
 - **RB03:** vhacdwwbs03.vha.med.va.gov

VINCI Use



Home About VINCI Skip to Content

VINCI Central

Welcome to the VA Informatics and Computing Infrastructure

Keyword Search Search

Laura Graham
VINCI Email: VINCI@VA.gov
Concierge Desk: 1-801-872-3324
Request Help Online

- Launch Workspace
- New To VINCI
- My VINCI Dashboard
- VINCI University
- VINCI Tube
- VINCIpedia/FAQ
- User Guides
- VINCI Services
- Computing Cloud
- Data Sources
- Applications
- Events
- Policies and Forms
- Quick Links

VINCI Happy Hour open forum Wednesday May 17th at 3pm ET

New To VINCI
Learn more about VINCI, VINCI Studies and the virtual environment.

My VINCI Dashboard
All your VINCI resources in one place!

VINCI University
Here you learn more

VINCI Tube
Videos on all things VINCI

VINCIpedia
VINCI's forum for research &

User Guides
Let us guide you through VINCI

VINCI Services

Computing Cloud

Data Sources

Applications

VINCI User Guides

Home About VINCI Skip to Content

VINCI Central

Welcome to the VA Informatics and Computing Infrastructure

Keyword Search Search

- Launch Workspace
- New To VINCI
- My VINCI Dashboard
- VINCI University
- VINCI Tube
- VINCIpedia/FAQ
- User Guides
- VINCI Services
- Computing Cloud
- Data Sources
- Applications
- Events
- Policies and Forms
- Quick Links

User Guides

- VINCI User Guides
- VINCI Software Guides
- OMOP Guides
- SAS Guides

Home About VINCI Skip to Content

VINCI Central

Welcome to the VA Informatics and Computing Infrastructure

Keyword Search

Laura Graham
VINCI Email: VINCI@VA.gov
Concierge Desk: 1-801-872-3324
[Request Help Online](#)

- Launch Workspace
- New To VINCI
- My VINCI Dashboard
- VINCI University
- VINCI Tube
- VINCIPedia/FAQ
- User Guides
- VINCI Services
- Computing Cloud
- Data Sources
- Applications
- Events
- Policies and Forms
- Quick Links

VINCI Happy Hour open forum Wednesday May 17th at 3pm ET

New To VINCI
Learn more about VINCI, VINCI Studies and the virtual environment.

My VINCI Dashboard
All your VINCI resources in one place!

VINCI University
Here you learn more

VINCI Tube
Videos on all things VINCI

VINCIPedia
VINCI's forum for research &

User Guides
Let us guide you through VINCI

VINCI Services

Computing Cloud

Data Sources

Applications

VINCI Standard Workspace Applications

[Standard Workspace](#) | [Development Workspace](#) | [Request a Workspace](#) | [Help](#)

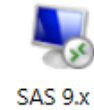
Important Instructions
Click once on any icon below to start the application.
 When prompted to enter your login, please enter it in the form of **DomainName\Login**.
 This site is only supported on the Internet Explorer browser with Java Script enabled.
 If you are not able to see the application icons below, [click here](#) to access the VINCI Standard Workspace Desktop to access all applications.

Current folder: /

7-Zip File Manager	Adobe Reader XI	ArcGIS Explorer	ArcMap 10.5	ArcReader 10.5	explorer	MATLAB R2016a	Microsoft Access 2010	Microsoft Excel 2010	Microsoft InfoPath Designer 2010
Microsoft InfoPath Filler 2010	Microsoft PowerPoint 2010	Microsoft Word 2010	Notepad++	NVivo 8	Protege 3.3.1	Protege 3.5	PSPP	PUTTY-CAC	R x64 3.3.1
RStudio	SAS 9.x	SAS Universal Viewer	SPSS Statistics & Modeler	SQL Server 2014 Management	StataMP 14 (64-bit)	StatTransfer 13	windiff	SAS Grid Tools	

VINCI Standard Workspace Applications

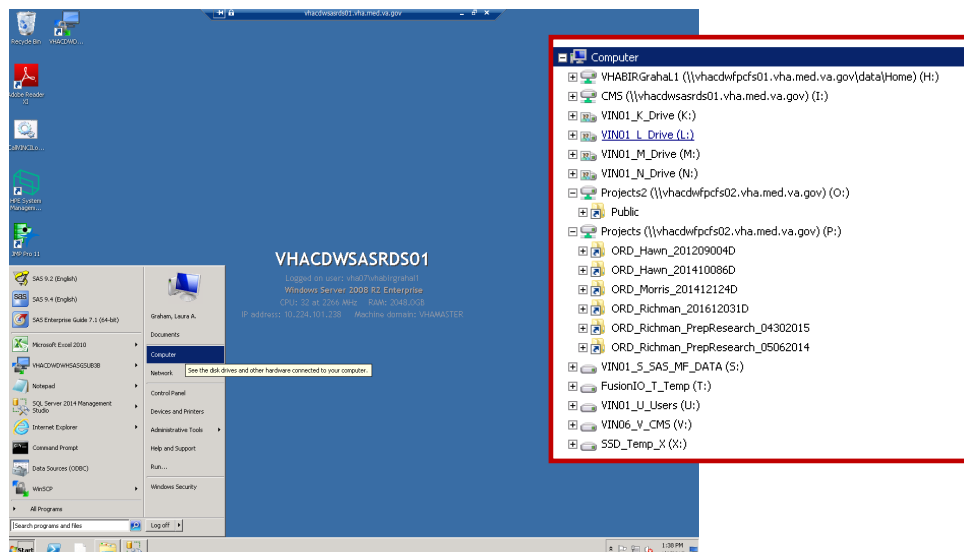
- Icons on the Standard Workspace link to remote desktop servers
- Your log-in and password are the same as for your VA computer (i.e. vhapal ...)
- The overall desktop application has limited software availability.
 - Log into each application separately
 - Each will be a separate desktop / server
- The SAS 9.x icon is a good place to start.



VINCI Folder Set Up

- Same as on a desktop
 - Hard Disk Drives (K, L, M, N, T, S, U, V, X)
 - Network Drives
 - P:/ drive (Projects)
 - O:/ drive (Projects2)
 - H:/ drive
- All project folders are typically found in the P:/ drive
- H:/ is your personal folder
- All project folders have a sub-folder named **Uploads** where uploaded information is included

VINCI Folder Set Up



VINCI – More Information

- More information on accessing and using the VINCI Workspace are found in the [VINCI Workspace User Guide](#)

VA Informatics and Computing Infrastructure	CONTROLLED DOCUMENT
VINCI Workspace User Guide	DOC. NO. 5002.20



SQL Server Management Studio

View Your Data

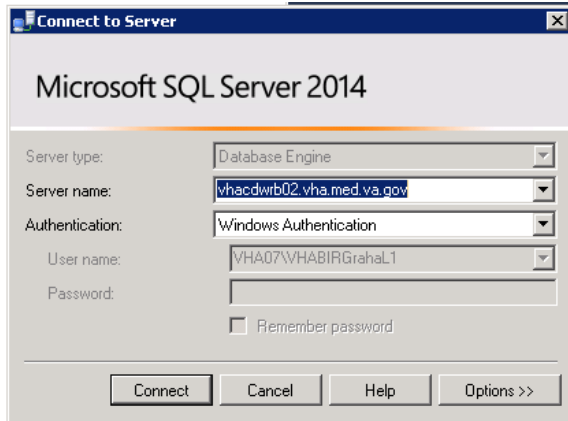
SQL Server Management Studio

- 2017 version
- Method for accessing and **viewing CDW domain** tables with SQL code
- Can also be used to **pull data in a .txt format** if needed
- Requires some initial set up to link to servers and some SQL knowledge but also has **point-and-click options** that make it easy to use

- The **best place to start looking** at data you've been given access to for your study

SQL Server Management Studio Set Up

- [VINCI Database User Guide](#)
 - Section 4.2.1 (pages 3-4) provides detailed information on SQL server set up

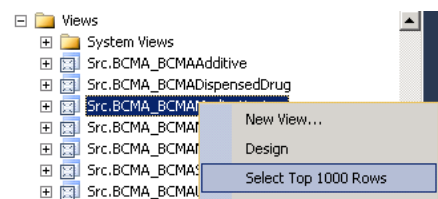


Default Settings To Use

Server Type: Database Engine
 Server Name: <Database Server Name>
 .vha.med.va.gov
 Authentication: Windows Authentication

SQL Server Management Studio Data Access

- [VINCI Database User Guide](#)
 - Section 4.2.2 (pages 4-7) provides detailed information on accessing databases
- Overview
 - **Databases > CDWork** contains your dimension tables
 - **Databases > ORD_XXXX > Views** contains your SQL data tables
 - Right-click and **“Select Top 1000 Rows”** for easy viewing of a data source



SQL Server Management Studio Data Pulls

- Not as efficient as a SAS pull but can be done
- First, set up your query. Your initial runs will be output to Grid.
- Once your query is set up
 - Use “**Query > Results To**” to change to “**Results to File**”
 - Then “**Query > Query Options...**” to set up your file but navigating through “**Results > Text**”

Applications for Analysis

Analyze Your Data

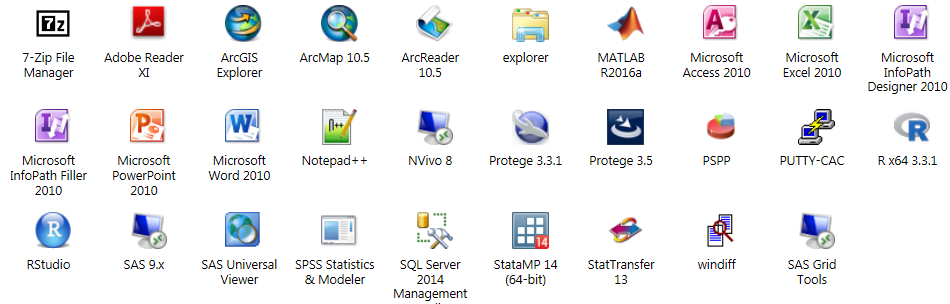
VINCI Standard Workspace Applications

Standard Workspace | Development Workspace Request a Workspace | Help

Important Instructions

Click once on any icon below to start the application.
 When prompted to enter your login, please enter it in the form of **DomainName\Login**.
 This site is only supported on the Internet Explorer browser with Java Script enabled.
 If you are not able to see the application icons below, [click here](#) to access the VINCI Standard Workspace Desktop to access all applications.

Current folder: /



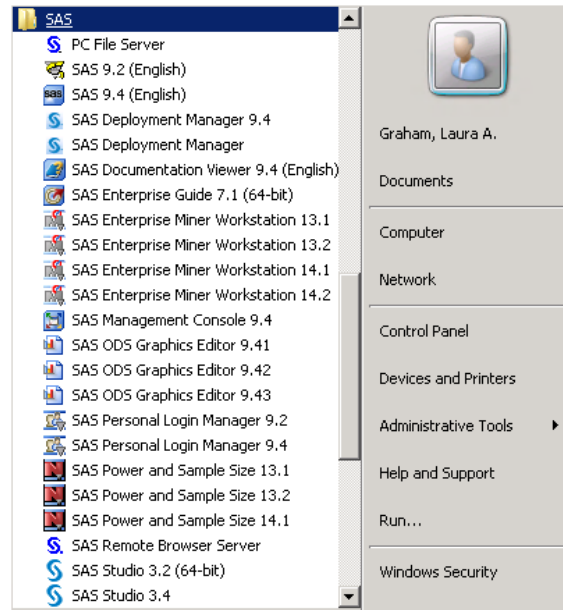
VINCI SAS

- Lots of **User Guides** for reference
- This course is only an intro but these will answer your more advanced questions.

SAS Guides			
Type	Guide Name	Description	Last Updated
	Grid - Where to Begin?	This document gets you started by providing information on using the Grid for new users.	2016-02-28
	Grid Computing in SAS 9.4	This guide is written by SAS on how to use the 9.4 SAS Grid.	2014-07-29
	How to use PCFILES engine	Demonstrates ways for the Grid platform to work with MS Excel, MS Access and STATA data files existing on Windows network shares.	2017-04-25

VINCI SAS

- Several versions of SAS are available for use.
 - SAS 9.2
 - SAS 9.4
 - SAS Enterprise Guide 7.1
- All offer different capabilities so check with all of them



VINCI SAS

- “Grid” refers to the Linux operating system where **SAS Enterprise Guide 7.1** runs
 - You will need a slightly different libname reference:
“/data/dart/2014/ORD_Smith_201401001D/”
 - You will also need to use WINS CP for data transfer and a PIV-exemption to access the Grid
 - Check out “[Grid – Where to Begin?](#)” for more information
- **SAS 9.2** and **9.4** are not on the grid and sometimes easier to work from.
 - Plus, SAS 9.2 is the only program with PROC TRAJ!

VINCI SAS Data Pulls

- I tend to use SAS 9.2 or SAS 9.4 to tunnel into the CDW and pull my data directly into SAS.
- This requires the following code:

```
proc setinit noalias; run;

* Source Libname;
libname SRC oledb provider=SQLNCLI11 datasource="VHACDWRB02.VHA.MED.VA.GOV"
properties=('initial catalog'=ORD_Morris_201412124D 'Integrated Security'=SSPI) schema=src;

* Dimension Tables;
libname DIM oledb provider=SQLNCLI11 datasource="VHACDWRB02.VHA.MED.VA.GOV"
properties=('initial catalog'=CDWWork 'Integrated Security'=SSPI) schema=Dim;
```

- The **Catalog** is changed to reference your project folder.
- “Datasource” may also need to be changed based on your project server.
- Schema will change based on what view you are looking at (SRC, Dim, ...)

VINCI SAS Data Pulls

- **PROC SQL** is typically how I pull and merge the data in SAS 9.4
- I initially set up the **test data pulls in SQL Management Studio** using a pull of the top 100 to test the queries
- Once my SQL code looks good then I **move it to SAS**, remove the brackets, update the libnames and run.

```
/****** Script for SelectTopNRows command from SSMS *****/
SELECT TOP 100 [Sta3n],[PatientSID],[AdmitDateTime],[DischargeFromFacility],
[DischargeDateTime],[PrincipalDiagnosisICD9SID]
FROM [ORD_Morris_201412124D].[Src].[Inpat_Inpatient]

PROC SQL;
CREATE TABLE new as
SELECT a.Sta3n,a.PatientSID,a.AdmitDateTime,a.DischargeFromFacility,
a.DischargeDateTime,a.PrincipalDiagnosisICD9SID
FROM Dim.Inpat_Inpatient as a;
Quit;
```

VINCI SAS Data Pulls

- This can get complicated with **multiple joins** and **Where statements**.

```

* Pulling Gabapentin;
PROC SQL;
  CREATE TABLE isqr.Gaba as
  SELECT
    x.scrssn, y.*, b.PatientSID, b.Sta3n, b.ActionDateTime, a.DrugClass,
    a.NationalDrugNameWithDose, a.NDC, a.DrugNameWithoutDose, a.LocalDrugMessage,
    a.DispenseUnit, a.DispenseUnitsPerOrderUnitNumeric,
    b.DosesGiven, b.DosesOrdered, b.UnitOfAdministration, c.ActionStatus,
    c.OrderSchedule
  FROM Dim.LocalDrug as a
  JOIN Src.BCMA_BCMADispensedDrug as b
    ON a.Sta3n = b.Sta3n AND a.LocalDrugSID = b.LocalDrugSID
  JOIN Src.BCMA_BCMAMedicationLog as c
    ON b.BCMA_MedicationLogSID = c.BCMA_MedicationLogSID AND b.Sta3n = c.Sta3n

  JOIN isqr.crosswalk2 as x
    ON x.PatientSID = c.PatientSID AND x.sta3n = c.sta3n
  JOIN base as y
    ON x.scrssn = y.scrssn AND
    (y.admtdt <= datepart(b.ActionDateTime) < y.hdisdt)
  WHERE a.NationalDrugNameWithDose LIKE '%GABA%' OR a.NationalDrugNameWithDose LIKE '%KETAMINE%';
QUIT;

```

VINCI SAS Data Pulls

- SQL pass-through language is also a quick way of tunneling into the SQL tables and pulling out data.

```

proc sql;
  connect to oledb as CDW(datasource="VHACDWRB02.VHA.MED.VA.GOV" provider=SQLNCLI11
    properties=('initial catalog'=ORD_Morris_201412124D 'Integrated Security'=SSPI));
  execute(
    select a.[sta3n], a.[patientsid], a.[admitdatetime], a.[dischargedatetime], a.[icd9sid], a.[icd10sid],
    d.[icd9code], b.[icd10code], e.[ICD9Description], c.[ICD10Description]
    into #in1
    from [Src].[Inpat_InpatientDischargeDiagnosis] as a
    left join [CDWWork].[Dim].[icd10] as b
      on a.[sta3n] = b.[sta3n] and a.[icd10sid] = b.[icd10sid]
    left join [CDWWork].[Dim].[ICD10DescriptionVersion] as d
      on a.[sta3n] = d.[sta3n] and a.[icd10sid] = d.[icd10sid]
    left join [CDWWork].[Dim].[icd9] as e
      on a.[sta3n] = e.[sta3n] and a.[icd9sid] = e.[icd9sid]
    left join [CDWWork].[Dim].[ICD9DescriptionVersion] as c
      on a.[sta3n] = c.[sta3n] and a.[icd9sid] = c.[icd9sid]
  ) by cdw;
  execute(

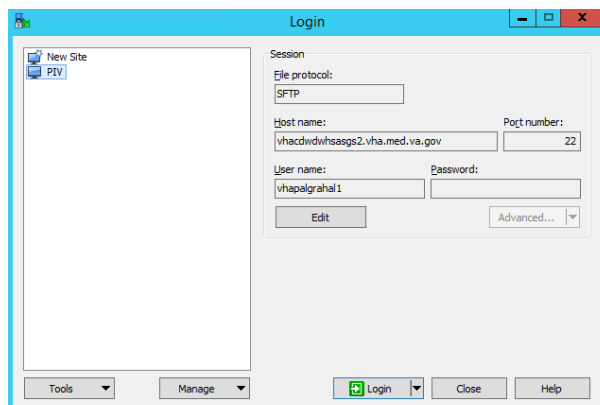
```

VINCI Enterprise Guide 7.1

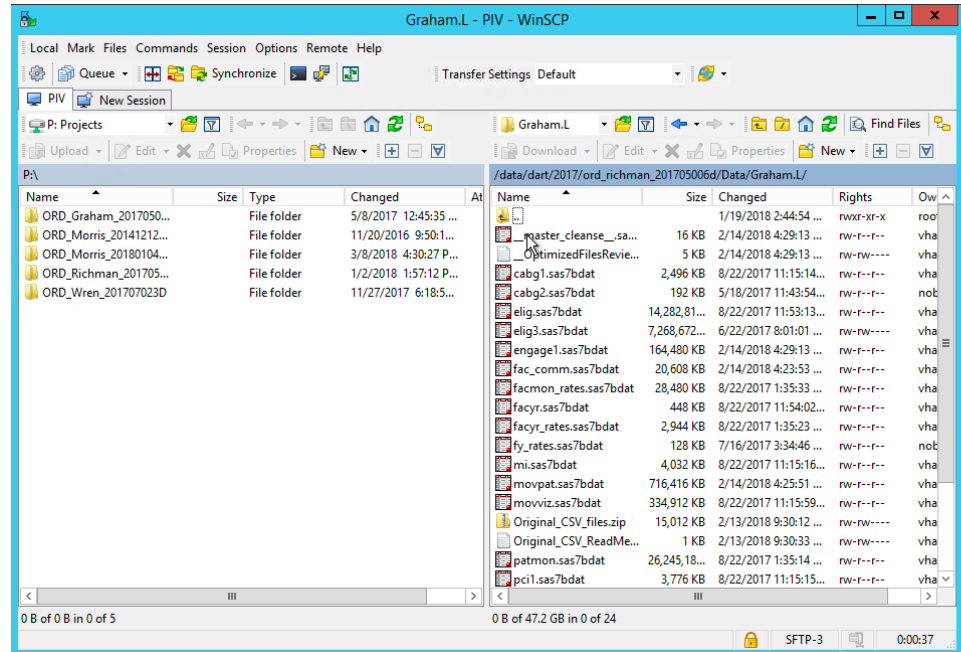
- VINCI SAS 9.4 and 9.2 are limited in the types of analyses that can be done
- VINCI Enterprise Guide offers everything and is by far the most capable of the SAS products but you have to set up links to the Grid and move your files to the Linux system
- I tend to do most of my data management and cleaning in 9.4 and then switch the analytic dataset to Linux for my analyses

VINCI Enterprise Guide 7.1

- The first step to using Enterprise Guide 7.1 is to check your Linux system registration by setting up **WinSCP**
- Once set up WinSCP is your method for moving files from PC to Linux
- This can be done by following the instructions in Using [Enterprise Guide with the Grid](#)

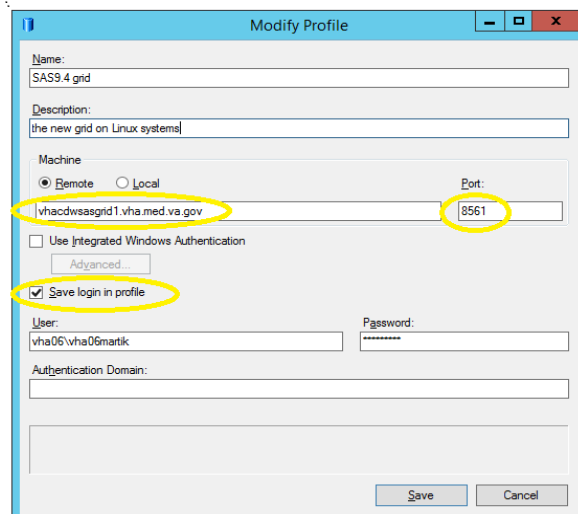


WinSCP



VINCI Enterprise Guide 7.1

- To use **SAS Enterprise Guide 7.1** you will need to set up your connection
- This can be done by following the instructions in Using [Enterprise Guide with the Grid](#)



Policies

Protect Your Data

Policies for Electronic Data

- **VA Handbook 6500**
 - Definitions and guidance for the VHA
- Key Points
 - **USB drives are forbidden**
 - Do not store sensitive information on a Hard Drive (C:/ Drive). Try not to store anything on a hard drive.
 - Sensitive information can never be transported, accessed, processed or otherwise used offsite. **DON'T EMAIL DATA!**

Types of Data

1. Individually Identifiable Health Information
 2. Protected Health Information (PHI)
 3. Personally Identifiable Information (PII)
 4. VA Sensitive Information/Data
- [VHA Handbook 6500](#) provides guidance on definitions and management

Disposal of Data

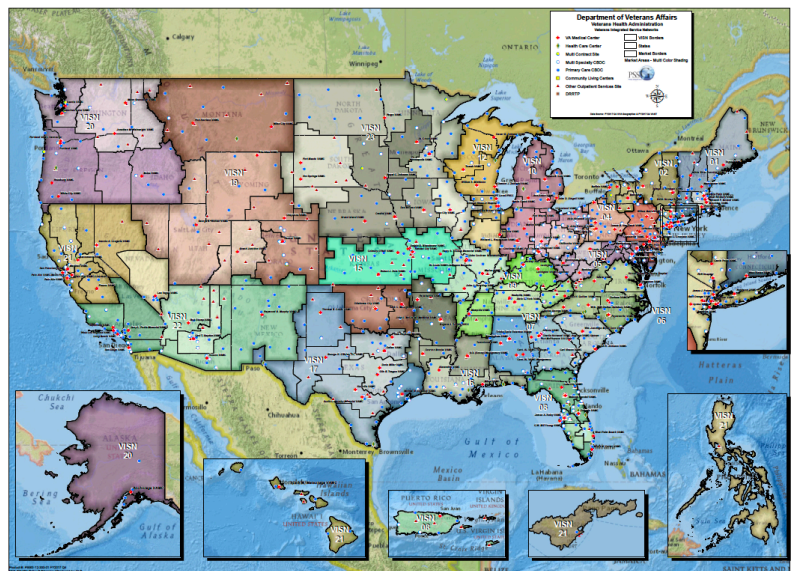
- Governed by the **VHA Record Control Schedule** ([RCS 10-1](#))
- Current guidance (Item 8300.6, Page III-8-6)
 - Temporary; cutoff at the end of the fiscal year after completion of the research project.
 - Destroy **6 years after cutoff**, may retain longer if required by other Federal regulations. (DAA-0015-2015-0004, item 0032)
 - If the investigator leaves VA, all research records are **retained by the VA facility where the research was conducted**.
 - If the grant is ongoing and the investigator leaves one VA facility to go to another VA facility, the investigator must obtain approval for a copy of relevant materials to be provided to the new VA facility's research office.
- **The investigator is not the grantee, nor does the investigator own the data.**



Fun Things to Do

Full Organizational View

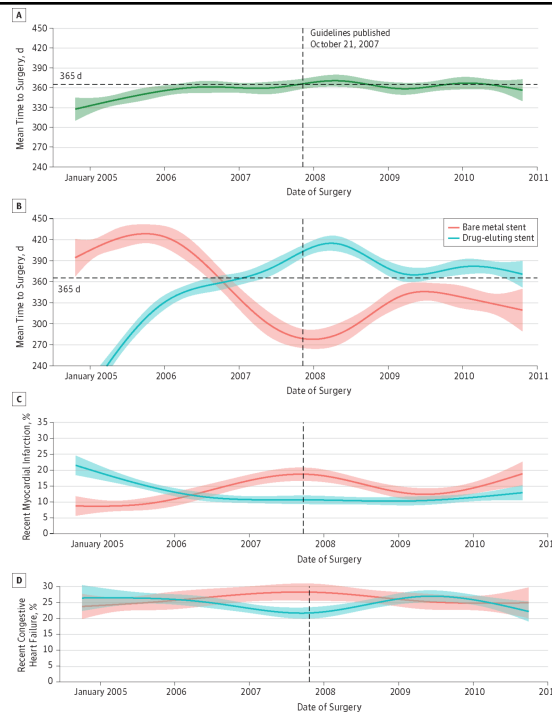
- >9 Million Enrolled Patients
- 170 VA Medical Centers (141 with Surgical Capacity)
- 1,063 Outpatient Centers



Longitudinal Data

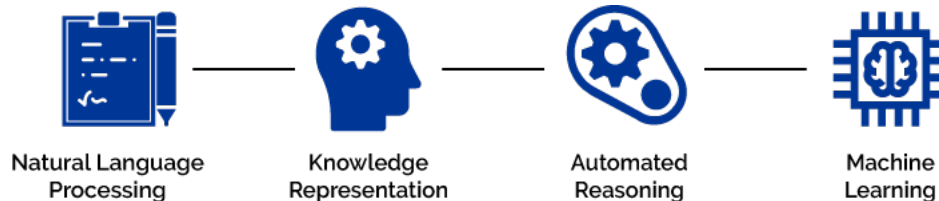
- All vitals and laboratory values collected in the VA are easily accessible
- Pharmacy prescriptions – inpatient, outpatient, and even evidence of non-VA prescriptions are available

JAMA Surg. 2014;149(11):1113-1120.
doi:10.1001/jamasurg.2014.2044



Natural Language Processing Capabilities

- Linguistic-computer science research
- Ability of a computer program to understand human language as it is written in free-text of clinical notes
- Several VA tools to work with



Thank you!