

Gator Vault Training

© 2012 Gator Vault, Inc. All rights reserved.



Introduction

2

Research computing often involves access to, and the use of, restricted data.

The University provides training to all staff, faculty, and researchers that have access to this broad category of information that requires special protections.

Let's begin by reviewing the definition of restricted data.

What is Restricted Data?

Data in any format that is collected, created, or maintained by or on behalf of the University and its affiliates, or within the scope of University activities

AND

subject to specific protections under federal or state law or regulations, or legal contracts.



Examples of Restricted Data

- ▶ Social Security numbers
- ▶ Credit card numbers
- ▶ Bank account and other financial numbers
- ▶ Florida driver's license number
- ▶ Medical records, account numbers, and photographs
- ▶ Student records such as grades, tuition and payments, and disciplinary status



See <http://privacy.ufl.edu> for a complete list.

Why Protect Restricted Data?

It's the law.

- ▶ HIPAA protects the privacy of health information
- ▶ FERPA protects student records
- ▶ Florida laws require the protection of health data
- ▶ Graham-Leach-Bliley regulations protect financial information
- ▶ PCI DSS is an industry standard that protects credit card information.

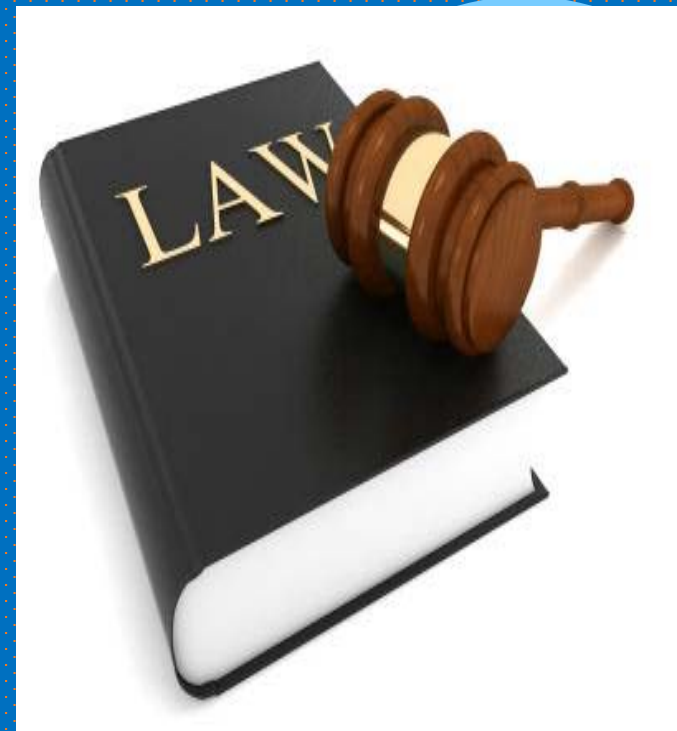
Our discussion will focus on HIPAA and Florida laws.



A Closer Look at HIPAA

State and federal laws protect restricted data residing in GatorVault.

HIPAA protects a specific type of health information; HIPAA regulations call this health data “protected health information”, or PHI.



A Closer Look at HIPAA

The Office of Civil Rights defines protected health information as:

- ▶ Individually identifiable health information that is transmitted or maintained in any form or medium by a Covered Entity or its Business Associate.

What is individually identifiable health information?

- ▶ Health information, including demographic information that
- ▶ Relates to an individual's physical or mental health or the provision of or payment for health care
- ▶ That directly identifies the individual, or could reasonably be used to identify the individual.

A Closer Look at HIPAA

Any information that links an individual with their physical or mental health condition such as:

- ▶ Person's name or the name of a relative
- ▶ Any address smaller than a state
- ▶ Dates such as birth & death, admission or discharge, date of service
- ▶ Medical record, account, & health plan beneficiary number
- ▶ SSN, email address
- ▶ Full face photograph

Even a single data element, such as date of birth, qualifies as PHI!

Florida Law

Florida Statute 501.171 protects Personal Identification Information (PII), which is defined as:

An individual's first name or first initial and last name in combination with one or more of the following data elements for that individual:

- ▶ A Social Security Number
- ▶ Driver's license or identification card number or similar government-issued ID
- ▶ Financial account number or credit or debit card number, and an access code that permits access to an individual's financial account.

Florida Law

10

Florida Statute 119.071 regulates the use of Social Security numbers in state agencies. The Privacy Office is responsible for establishing University policies that align with this state law.

A key policy established for the protection of SSNs is the requirement to obtain WRITTEN permission from the Privacy Office to use, disclose, and/or store SSNs for purposes not previously approved.

The Privacy Office website contains a list of approved uses and disclosures. <http://www.privacy.ufl.edu>.

Non-Compliance

11

Key Concept: All users are responsible for the protection and appropriate use of restricted data.

There are severe consequences for the University and for individuals if the privacy or security of restricted data is violated.

Florida law states that "Any person who willfully and without authorization fraudulently uses ... personal identification information without first obtaining that individual's consent, commits... a felony."

Non-Compliance

12

Federal sanctions for an individual responsible for a privacy incident include:

- ▶ Jail time
- ▶ Fines
- ▶ Civil “invasion of privacy” litigation by affected individuals is also a possibility.

Employee sanctions at UF range from a verbal warning up to and including termination.

Non-Compliance

13

The University assumes direct and indirect costs for non-compliance:

- ▶ Loss of the public's trust
- ▶ Damages to UF's reputation
- ▶ Regulatory fines and penalties
- ▶ Expenses associated with notifying patients of a breach

HIPAA and Florida law have breach notification requirements.

Breach Notification

14

UF defines a data breach as “an unauthorized access, acquisition, use, or disclosure of restricted data”.

In practical terms, this means that UF must notify patients- in writing- when their information has been compromised.



Breach Notification

15

Key Concept:

Data users must recognize a potential privacy or security incident and report it immediately to the Privacy Office.

The Privacy Office will conduct an investigation and, when required by HIPAA or Florida law, notify the affected patients.

Recognizing a Privacy Incident

16

Again, an incident is an inappropriate use or disclosure of protected health information. Examples include:

- ▶ Lost or stolen mobile device
- ▶ Sharing PHI with another GatorVault user who does not have IRB authorization for that PHI.
- ▶ Emailing unencrypted PHI to someone outside the UF domain



Reporting a Privacy Incident

17

Known or suspected incidents must be reported immediately to your supervisor or the Privacy Office:

- ▶ Privacy Office Direct: (352) 273-1212
- ▶ Privacy Hotline: (866) 876-4472
- ▶ Email: privacy@ufl.edu



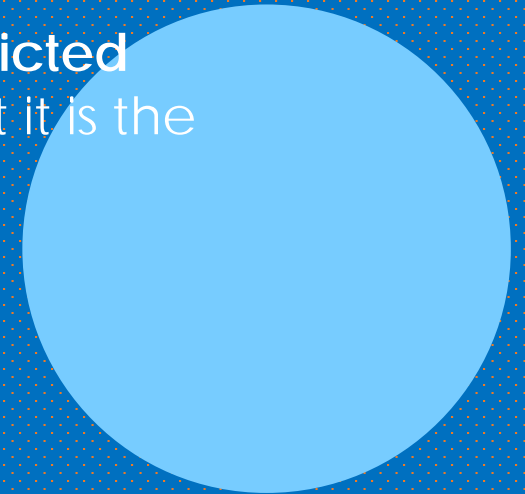
Failure to report will result in disciplinary action.

Privacy Wrap-Up

18

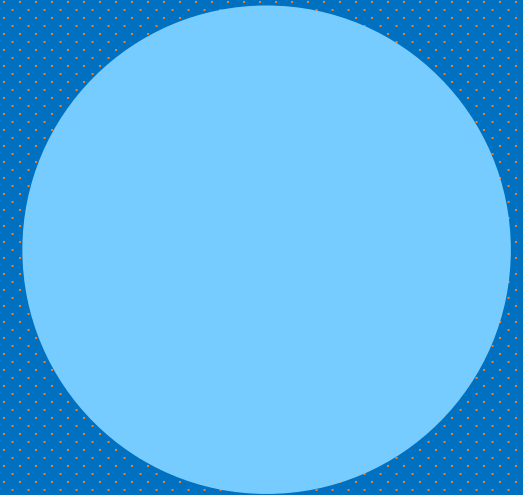
The University has a legal obligation to protect **restricted data**. But perhaps the most important reason is that it is the right thing to do.

Now we'll discuss GatorVault.



What is GatorVault?

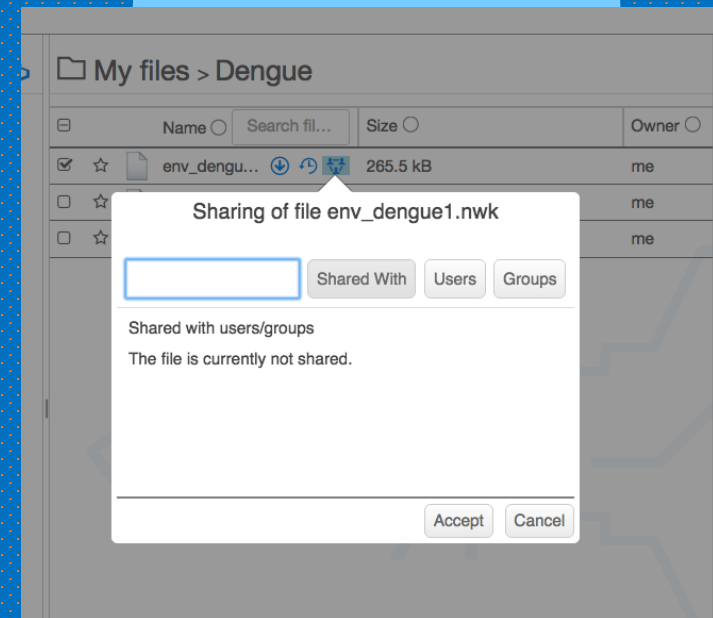
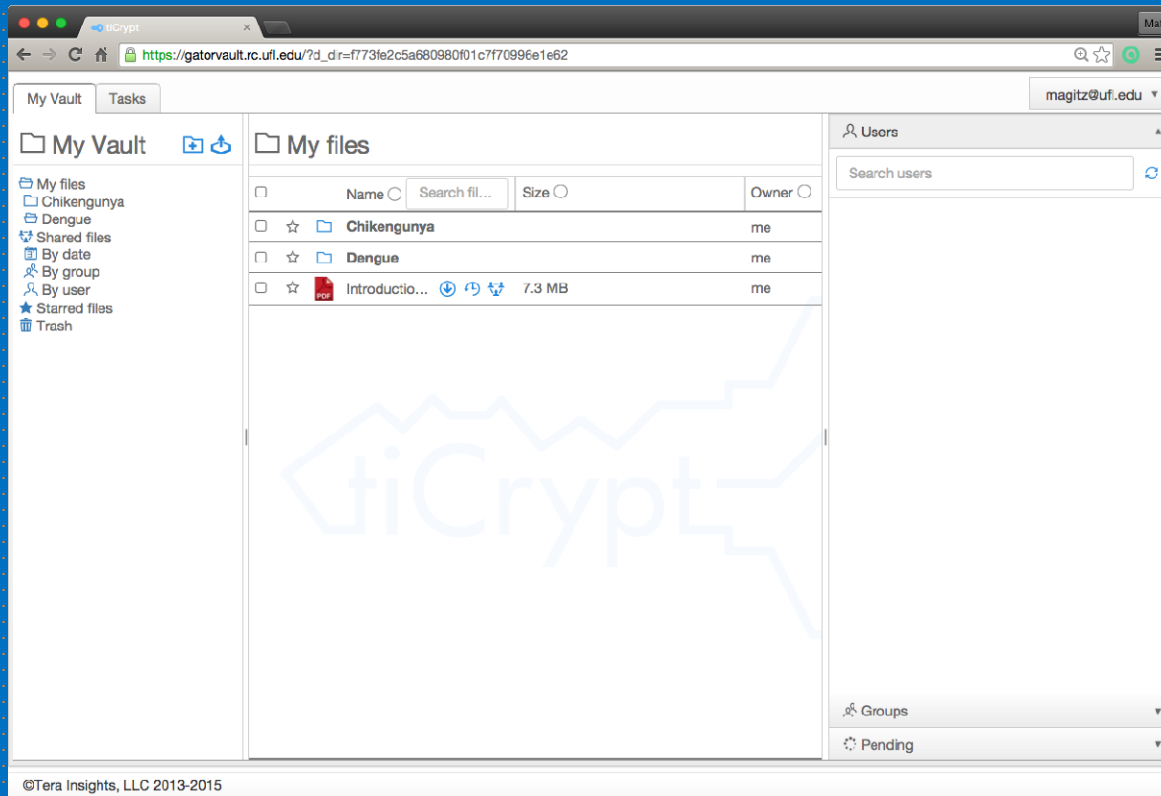
- ▶ Secure Data Storage and Collaboration
 - ▶ GatorVault allows storing and collaborating on data
 - ▶ Including PHI and restricted data
 - ▶ Securely share with authorized collaborators
- ▶ Secure Data analysis
 - ▶ GatorVault brings your data to robust computing resources for analysis
 - ▶ Virtualized computing resources allow analysis of data within the GatorVault environment
 - ▶ Windows
 - ▶ Linux



What is GatorVault?

20

► Secure Data Storage and Collaboration



What is GatorVault?

21

► Secure Data analysis

The screenshot displays the RStudio environment. The left pane shows a list of demos with descriptions:

- graphics: A show of some of R's graphics capabilities
- image: The image-like graphics builtins of R
- persp: Extended persp() examples
- plotmath: Examples of the use of mathematics annotation

Demos in package 'grDevices':

- colors: A show of R's predefined colors()
- hclColors: Exploration of hcl() space

Demos in package 'stats':

- glm.vr: Some glm() examples from V&R with several predictors
- lm.glm: Some linear and generalized linear modelling examples from 'An Introduction to Statistical Modelling' by Annette Dobson
- nlm: Nonlinear least-squares using nlm()
- smooth: 'Visualize' steps in Tukey's smoothers

Use 'demo(package = ,packages(all,available = TRUE))' to list the demos in all *available* packages.

The console shows the following output:

```
***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 10.452  on 5  degrees of freedom
Residual deviance: 5.173  on 4  degrees of freedom
AIC: 38.741

Number of Fisher Scoring iterations: 3

> op <- par(mfrow = c(2,2), oma = c(0,0,1,0))
> plot(glm.p84) # well ?
Hit <Return> to see next plot: |
```

The Environment pane shows the following objects:

Object	Type	Value
weight	num [1:30]	4.17 5.58 5.18 6.11 4.5 4.61 5.17 4.53 5.33 5...
x	num [1:6]	55 52 57 55 50 50
z	List of 30	
z.04	List of 30	
z0	List of 30	
z1	List of 30	
z2	List of 30	
z1	List of 30	
z2	List of 30	

The bottom pane displays a plot titled "Dobson's Birth Weight Data". The x-axis is labeled "age" and ranges from 35 to 42. The y-axis is labeled "birthw" and ranges from 2400 to 3200. The plot shows two data series: "M" (males, represented by black crosses) and "F" (females, represented by red inverted triangles). Both series show an increasing trend in birth weight with age, with males generally having higher birth weights than females. Two fitted curves are shown: a black line for males and a red line for females.

What is GatorVault?

- ▶ No need to move data from system to system
- ▶ Scalable computing capacity with the security needed for protected data

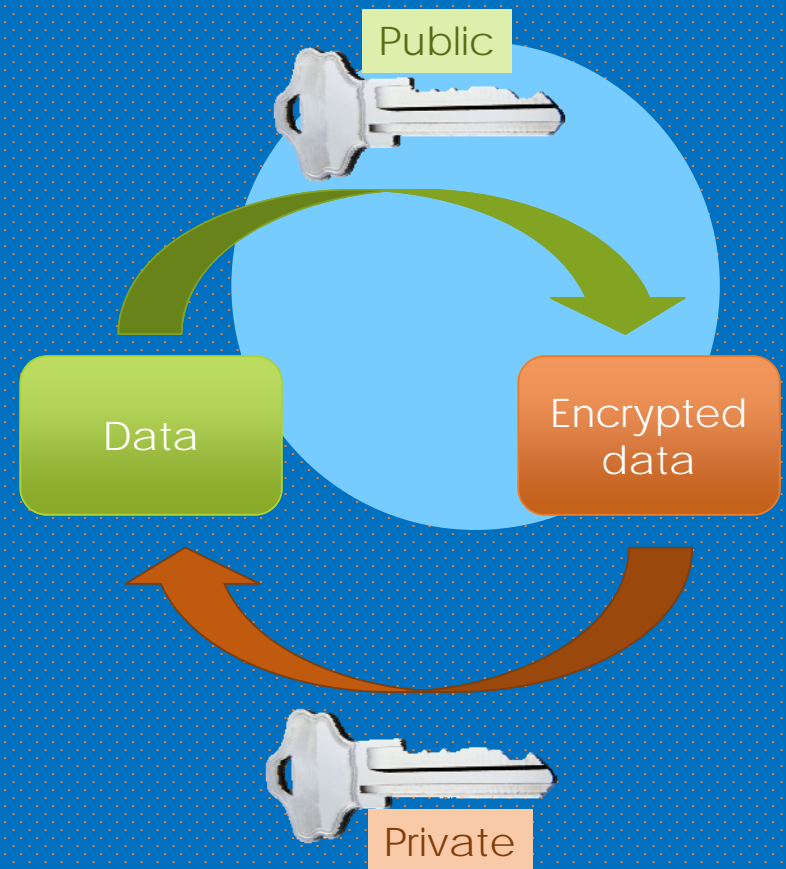
Secure
Data Storage
Collaboration
and
Analysis



Public-Private Key Basics

23

- ▶ GatorVault works by encrypting all data using public-private key-pairs
- ▶ The only way to decrypt the data is with the **private key**
 - ▶ You are the only one with the private key
 - ▶ **If you lose the private key, the data cannot be decrypted**—even system administrators cannot view your data!
 - ▶ Backup your private key
 - ▶ Never share your private key
 - ▶ Private key is password protected



Use of GatorVault

- ▶ Eligibility:
 - ▶ UF Faculty working with protected data
 - ▶ Their students and staff
 - ▶ Their authorized collaborators

GatorVault

PHI

Securely store and analyze patient data without having to de-identify

Intellectual Property

Securely manage sensitive IP data

ITAR

Securely work with restricted data and applications

Cost of GatorVault

25

- ▶ Interested faculty should contact Research Computing to discuss needs, costs and review procedures
 - ▶ For current prices see <http://www.rc.ufl.edu/services/rates/>



The screenshot shows the UF Research Computing website's Price Sheet page. The page is titled "Price Sheet" and includes a navigation menu with options like Home, Services, and Price Sheet. A sidebar on the left lists various services such as Computation, Storage, and Software Consulting. The main content area displays a table with columns for Service, Description, and Price. The table is divided into two sections: Computation Services and Storage Services. A note at the top of the table states: "Please note: we are in the process of updating our website so all information is subject to change."

Service	Description	Price
Computation Services		
Compute nodes - investor	Priority access to the number of NCUs invested in for five years	\$200 / NCU
Compute nodes - Per Hour	Computing is charged monthly for NCU hours used	\$0.04 / NCU-hour
NVIDIA GPUs	Access to the number of NVIDIA K80 NCU accelerators invested in for five years	\$500 / NCU
Intel Xeon Phi	Access Intel Xeon Phi E100p accelerators	No cost
Storage Services		
Lamp-term	Storage for data which is no longer needed for calculations, but for which some means of convenient access is required. No data protection beyond the basic RAID mechanisms is provided.	\$125 / TB / year
Replicated long-term	Storage for data which is no longer needed for calculations, but for which some means of convenient access is required. A replica on another disk system in another building is included for the safety of the data.	\$250 / TB / year
Tape backup	Enterprise level tape backup is provided by EIO NSAM services	\$80 / TB / year \$100 / TB / installer

Getting an account

26

- ▶ Accounts are created at:
 - ▶ <https://gatorvault.rc.ufl.edu/register.html>
 - ▶ Fill out the form
 - ▶ Backup your private key
- ▶ The account needs to be activated by a Research Computing staff member



GatorVault Registration

1. Generating public-private keypair

For secure access a public/private keypair is generated in your browser. This may take a few moments.

The key-pair has been generated.

2. Encrypt the private key with a password

The public key is stored on the server, while the private key is encrypted with your password and stored on the browser's local storage. The password is not stored anywhere (and cannot be reset). Access to files will be lost if the private key or password is lost.

Username (your email address)

Email

First name

First name

Last name

Last name

Password

Password

Retype password

Confirm Password

I understand that I am responsible for protecting my private key

I read and agree with the terms of use

Register

3. Save your (encrypted) private key

Save your file with your encrypted private key to the 'downloads' folder, and move it to your USB stick. You will need the file with the private key to access GatorVault on another machine.

GatorVault Requirements

Before using the GatorVault system, UF staff and researchers must successfully complete the appropriate training module annually:

- ▶ HIPAA General Awareness
- OR
- ▶ HIPAA & Privacy for Researchers

Both courses are available at <https://mytraining.hr.ufl.edu>.

Your position may require completion of Training Module IRB802 -OLT IRB01 Mandatory Local Training Refresher.

GatorVault Access

28

- ▶ Connecting to GatorVault immediately leads to GatorLink authentication
- ▶ GatorLink account is not sufficient to obtain access
- ▶ Authorization is checked and granted by GatorVault



GatorVault Secure Key

- ▶ The first thing to do
- ▶ Create RSA-2048 public/private key pair
 - ▶ Used for all authentication, authorization, and logging in GatorVault
- ▶ Encrypt the private key with AES-256 using a password as seed
 - ▶ Save this in a file, remember the password
 - ▶ Strength of the password will be enforced
- ▶ Make a backup of the password protected private key
- ▶ GatorVault does not see or keep your private key
 - ▶ Lose the key
 - ▶ Lose your files



Questions or Concerns?

30

Contact the Privacy Office

Tigert Hall, Room G-24
352-273-1212
privacy@ufl.edu

Research Computing

Physics 2334
2001 Museum Rd.
support@rc.ufl.edu

Next Steps

31

Now that you have completed the GatorVault training module, complete the assessment.

