Florida State University Security Best Practices for Controlled-Access Data Subject to NIH Genomic Data Sharing (GDS) Policy

July 2016, Version 1.0

Introduction

This document is intended for Florida State University (FSU) investigators who are granted access under the NIH Genomic Data Sharing (GDS) Policy to controlled-access human genomic and phenotypic data that are maintained in NIH-designated data repositories. It provides an outline of the university and NIH's expectations for the management and protection of NIH controlled-access data transferred to and maintained by FSU whether in their own institutional data storage systems or in cloud computing systems. Although controlled-access data do not contain direct identifiers, the data are sensitive and must be protected. The principles governing access and use of such data are outlined in the GDS Policy and individual Data Use Certification (DUC) Agreements that FSU investigators submit as part of the process of requesting access to controlled access data. This process is intended to ensure that NIH controlled-access genomic and phenotypic data are kept secure and no one other than users approved by NIH is able to access the data.

NIH controlled-access human genomic and phenotypic data is classified as "Protected' under the *University Information Classification Guidelines*.

The information contained in this document is targeted at two audiences: FSU scientific professionals including institutional signing officials and investigators who will use the data, and information technology professionals and operations staff working for both central IT organizations and embedded within the responsible FSU research groups. Accordingly, this document is split into two main sections focused on each of these groups.

FSU Policies providing support for IT security and privacy controls include:

4-OP-H-5 Information Security Policy
4-OP-H-12 Information Privacy Policy

The controls in this document must be followed regardless of the location (local, central, or cloud computing devices) of the computing infrastructure used to store, process, or transmit the protected information.

Not all dbGaP studies require the submission of a System and Security Plan (SSP) for review. You should check the Data Use Certification (DUC) and application requirements for the studies of interest to determine if a SSP (See Appendix A) is required as part of your dbGaP data access request.

Information for Scientific and Administrative Staff

General Considerations

Under the GDS Policy, FSU is ultimately responsible for maintaining the confidentiality, integrity and availability of the data to which it is entrusted by the NIH. Failure to provide appropriate controls can subject investigators or FSU to sanctions defined by the GDS Policy. It is therefore essential that all recipients of

controlled access data understand their responsibilities for ensuring appropriate information security and privacy controls and that they work with their local and enterprise IT organizations to effectively implement those responsibilities.

Part of having an information security mindset is being aware of the multiple dimensions of ***access control*** and ***accountability*** at all times. This means ensuring that passwords and/or access devices smart cards, soft or physical tokens, etc.) are physically safe, strong and not shared with anyone and that data is both physically and logically (i.e. electronically) secure. Particular care must be taken with copies of data on portable electronic media and devices (e.g. laptops, tablets, USB thumb drives, tapes, etc.). Generally speaking, users should avoid putting controlled access data on such devices wherever possible. If it is necessary, such devices must be encrypted and should be treated as if they were cash, with appropriate physical and electronic controls, including remote wipe capability wherever possible. In addition, please remember that collaborators at different institutions must file a separate data access request even if they are working on the same project.

The *FSU Information Privacy Policy* covers research data privacy and security in the following paragraphs under Section C. Standards for Specific Information Types:

Research Information

University units conducting research must be aware of appropriate privacy restrictions for information transmitted, stored, or processed as part of research projects. Research projects are also a required component of a University unit's yearly data classification, risk assessment, and risk mitigation planning.

Legal privacy restrictions include, but are not limited to, the Health Insurance Portability and Accountability Act (HIPAA), International Traffic in Arms Regulations (ITAR), The Belmont Report (1979) and 2.1 Code of Federal Regulations Title 45 Part 46: The Common Rule concerning the protection of human subjects, other federal or state legal requirements, and contractual research information privacy restrictions. In addition, University units must protect the privacy of protected or private research information with appropriate information privacy and security controls such as those published by the National Institute of Standards and Technology (NIST), ISO, or Federal Information Security Management Act (FISMA). Required information privacy and security controls extend to any device used to transmit, store or process protected or private research information.

Finally, remember that data downloaded from NIH-designated data repositories must be destroyed when they are no longer needed or used, or if the project is to be terminated and closed-out in the dbGaP Authorized Access System. FSU recommends units' review *National Institute for Standards and Technology (NIST) Special Publication 800-88, Revision 1, Guidelines for Media Sanitization* for electronic data erasure and utilize cross-cut shredders for paper documentation. Investigators may retain only encrypted copies of the minimum data necessary to comply with institutional scientific data retention policy, and any data stored on temporary backup media as are required to maintain the integrity of the general institutional data protection (i.e. backup) program.

Confidentiality Statement and Privacy Training

Per *4-OP-H-12 Information Privacy Policy*, *Section B. Access and Use:*

Confidentiality Statement and Privacy Training

Signed Employee Statement of Understanding Regarding Confidentiality and training are required for FSU personnel with authorization to access or process protected or private information:

1. Each FSU position requiring access to protected or private information must be reflected in the position description.
2. For each person requiring access to protected or private information, signed Employee Statement of Understanding Regarding Confidentiality must be maintained on file unit and be available for audit. This information may be stored in a digital or paper format.
3. Employees designated as having access to select protected information (e.g., HIPAA) may be required to sign agreements acknowledging special confidentiality controls necessary to meet specific legal or contractual privacy requirements. These agreements are in addition to a signed FSU Employee Statement of Understanding Regarding Confidentiality document.
4. Each unit must train its employees on the requirements to safeguard protected or private information. This training should occur prior to employee access of protected or private information or as required by legislation or contractual obligation.
5. As verification of participation, each University unit must maintain rosters of participants in online or in-person privacy training in an electronic or paper format.

Additional Information Related to the Use of Cloud Computing

Cloud computing, as defined by the NIST, is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or cloud service provider interaction. In contrast to traditional computing on local servers and hardware, cloud computing often entails the transfer and storage of controlled-access data on systems managed by a third party. Cloud computing offers a number of advantages for authorized investigators but also requires additional security considerations.

Investigators who wish to use cloud computing for storage and analysis will need to indicate in their NIH Data Access Request (DAR) that they are requesting permission to use cloud computing and identify the cloud service provider or providers that will be employed. They also will need to describe how the cloud computing service will be used to carry out their proposed research. In addition, any FSU entity utilizing cloud based services to store NIH protected data should ensure that they understand the security policies and practices utilized and recommended by their cloud service provider of choice. Only FSU-procured commercially contracted cloud services should be used to store NIH protected data. The contract for this service must include the university's *IT Security and Privacy Terms and Conditions* as part of the procurement process. Because the use of cloud computing has the potential for being higher risk than using local infrastructure, NIH strongly recommends that you consult with the university's Director of IT Security and Privacy in addition to your unit's IT support staff to ensure that an appropriate security plan is developed and that necessary technical, training and policy specified controls are in place prior to the migration of protected information to cloud environments.

The *FSU 4-OP-H-12 Information Privacy Policy* stipulates the required process for contracting cloud based services under section B. Access and Use:

Third-party Access to Protected or Private Information

FSU may choose to contract with a third-party for the collection, storage, or processing of information, including protected or private information. The third-party may offer services in the form of hosting, outsourcing, or private/public cloud computing services.

If FSU decides to contract a third-party for the processing of protected or private information, this must be regulated in a written agreement, in which the rights and duties of FSU and the third-party contactor in addition to any subcontractors

engaged by the primary third-party contractor are specified. A third-party contractor shall be selected that will guarantee the technical and organizational security/privacy measures required in this privacy policy and provide sufficient guarantees with respect to the protection of the information.

FSU provides a terms and conditions document containing privacy and security provisions for information sharing agreements involving protected or private information.

Terms and Conditions document: Suggested Contractual Provisions for the External Sharing of FSU Information Classified as Protected or Private.

A third-party contractor should also be contractually obligated to process protected or private information only within the scope of the contract and the directions of FSU. Processing of protected or private information may not be undertaken for any other purpose.

Information for IT Professionals

Local Infrastructure Guidance

General Information Security Guidelines

- When using FSU or unit supported networks, make sure NIH protected files are never directly Internet assessable (with the exception of such connections as are required to download data from source repositories). Computing infrastructure should be behind local and/or ITS-CORE managed firewalls that block direct access to computing devices used to download, store, or process NIH protected information. For cloud infrastructure, investigators must restrict external access to instances and storage under the investigator's control.

- Data must never be posted on servers in any fashion that will make them publically accessible, such as an investigator's (or institution's) website, because the files can be discovered and indexed by Internet search engines such as Google or Bing.

- Units must not set up web or other electronic services that host data publicly, or that provide access to other individuals that are not listed on the Data Use Request even if those individuals have access to the same dbGaP data. Providing such access requires that an organization be an NIH Trusted Partner, with different requirements above and beyond those required for access to NIH controlled data.

- Utilize strong authentication technology for access control. Two factor authentication technologies (e.g., Duo) are preferred. When using single factor passwords for central computing resources, engage ITS-Identity Management to set policies that mandate the following requirements for enterprise logons. In addition, devices not using enterprise logon or those under the control of unit Active Directory (AD) authentication not part of the FSU master AD must adjust their authentications controls to meet the parameters below:

    - Minimum length of 12 characters
    - Does not contain user names, real names or company names
    - Does not contain a complete dictionary word
    - Contains characters from each of the following groups: lowercase letters, uppercase letters, numerals, and special characters
    - Passwords should expire every 120 days.

- Avoid allowing users to place controlled access data on mobile devices (e.g. laptops, smartphones, tablets, mp3 players) or removable media such as USB thumb drives (except where such media are used as backups and follow appropriate physical security controls). If data must be placed on mobile devices, it must be encrypted and require a passcode for devices when an available option for access. The FSU Information Security and Privacy Office (ISPO) and NIH recommend the use of the currently approved validated encryption technologies on the NIST Advanced Encryption Standard Algorithm Validation List.

- Keep all software patches up-to-date with critical patches applied to computing devices storing, transmitting, or processing NIH protected information. Critical patches must be applied within 30 days of a patch release. University units are responsible for using the ISPO offered Nexpose vulnerability

scanner to inventory vulnerabilities on devices every 30 days. Credentialed scans of devices must be conducted using Nexpose to fully discover the state of the operating systems and application supported on the device.

- Computing devices, including mobile devices, used to access, store, or process protected data must have a screen lock that requires a PIN or password to unlock after 30 minutes of inactivity.

Physical Security Guidelines

- Data that are in hard copy or reside on portable media, e.g., on a USB stick, CD, flash drive or laptop, should be treated with appropriate controls. Such media must be encrypted and stored in a secured in a locked facility with access granted to the minimum number of individuals required to efficiently carry out research.

- Restrict access to printers, plotters, and other devices used to display NIH protected information in a hard copy format. Limit physical access to printers or use one-time authentication codes for print jobs that include NIH protected information when the printer device cannot be physically secured from access by individuals not approved to view protected information.

- Restrict physical access to all servers, network hardware, storage arrays, firewalls and backup media by key or card lock to only those that are required for efficient operations.

- Log access to secure facilities such as network closets and server rooms, ideally with electronic authentication. However, paper or electronic logs are acceptable for audit purposes.

- Physical security controls for handling protected data are defined in *4-OP-H-12 Information Privacy Policy*:

  Physical Security Access Restrictions

  Offices and storage facilities that maintain protected or private information locally must:

  1. Ensure that all protected or private information in hardcopy or electronic form is secure in their work area at the end of the day and when they are expected to be gone for an extended period.
  2. Computer workstations processing, transmitting, or storing protected or private information must be secured by locked rooms when the workspace is unoccupied.
  3. Any protected or private information should be removed from the desk and locked in a drawer when the desk is unoccupied and at the end of the work day if the room cannot be secured.
  4. File cabinets containing protected or private information must be kept closed and locked when not in use or when not attended.
  5. Keys used for access to resources holding protected or private information must not be left at an unattended desk.
  6. Passwords may not be left on sticky notes posted on or under a computer, nor may they be left written down in an accessible location.
  7. Printouts containing protected or private information should be immediately removed from the printer in unsecured areas.
  8. Upon disposal, documents containing protected or private information should be shredded or placed in the lock confidential disposal bins. Electronic media containing protected or private information that is no longer needed should be physically destroyed (e.g., shred, degauss) or wiped by electronic methods to render the information unreadable and unrecoverable as stipulated in National Institute of Standards and Technology-Special Publication 800-88 Revision 1Guidelines for Media Sanitization.

9. Whiteboards containing protected or private information should be erased unless they are in secured areas. In addition, whiteboards with protected or private information should not be facing external windows unless blinds are drawn down to prevent unauthorized viewing of content.
10. Portable computing devices containing protected or private information such as laptops, phones, tablets, CDROMs, DVDs, USB flash drives should be secured in locked rooms, file cabinets, or locked drawers after normal work hours.

Controls for Servers

- Keep servers from being accessible directly from the Internet, (i.e. must be behind a firewall or not connected to a larger network) and disable unnecessary services. Use server and device hardening guides from the Center for Internet Security (CIS) or the *United States Government Configuration Baseline* (USGCB) to harden devices before putting them into a production environment.

- All servers must have a credentialed Nexpose vulnerability scan run against the device prior to putting the device into a production environment. All critical vulnerabilities or severe vulnerabilities with publically available exploit kits must be remediated prior to putting the server into the production environment.

- Enforce principle of least privilege to ensure that individuals and/or processes are granted only the rights and permissions necessary to perform their assigned tasks and functions, but no more.

- Secure controlled-access genomic and phenotypic data on systems from other users (restrict directory permissions to only the owner and approved research group members) and if exported via file sharing, ensure restricted access to remote systems.

- If accessing systems remotely, use encrypted data access (such as Secure Shell (SSH) or the university's Virtual Private Network (VPN)). It is preferred to use a tool such as Remote Desktop (RDP), X-windows or Virtual Network Computing (VNC) that does not permit copying of data and provides "View only" support.

- If data is used on multiple systems (such as a computer cluster), ensure that data access policies are retained throughout the processing of the data on all the systems. If data is cached on local systems, directory protection must be kept, and data must be removed when processing is complete.

- Requesting investigators must meet the spirit and intent of these protection requirements to ensure a secure environment 24 hours a day for the period of the NIH agreement.

Source Data and Control of Copies of Data

- Approved users must retain the original version of the encrypted data, track all copies or extracts and ensure that the information is not divulged to anyone except authorized staff members at FSU. Researchers must also control physical copies of data and provide appropriate logging on machines where electronic protected data is resident. Cataloging of electronic data is also mandated by the university's privacy policy.

- Collaborating investigators from other institutions must submit an independent DAR and be approved by NIH prior to accessing information restricted under a NIH agreement with FSU.

Destruction of Data

- Data downloaded from NIH-designated data repositories must be destroyed if they are no longer needed or used, or if the project is to be terminated and closed-out in the dbGaP Authorized Access System. Delete all data for the project from storage, virtual and physical machines, databases, and random access archives (i.e., archival technology that allows for deletion of specified records within the context of media containing multiple records). Destroy media according to (NIST) Consult NIST Special Publication (SP) 800-88, Revision 1, for approved media erasure guidelines for local media destruction.

- Investigators and Institutions may retain only encrypted copies of the minimum data necessary at their institution to comply with institutional scientific data retention policy and any data stored on temporary backup media as are required to maintain the integrity of the institution's data protection program. Ideally, the data will exist on backup media that is not used by other projects and can therefore be destroyed or erased without impacting other users/tenants. If retaining the data on separate backup media is not possible, as will be the case with many users, the media may be retained for the standard media retention period but may not be recovered for any purpose without a new Data Access Request approved by the NIH. Retained data should be deleted at the appropriate time, according to institutional policies.

- Shred hard copies and CD ROMs or other non-reusable physical media.

- Ensure that backups are reused (data overwritten) and any archive copies are also destroyed.

- Sanitize media according to (NIST) Consult *NIST Special Publication (SP) 800-88, Revision 1,* for approved media erasure guidelines for local media destruction. Ensure any NIH protected information hosted by vendors in the cloud use NIST approved standards for information destruction upon the conclusion of contracted services involving NIH protected information. Delete electronic files securely. For personal computers, the minimum would involve deleting files and emptying the recycle bin or equivalent with equivalent procedures for servers.

Additional Guidance for Cloud Computing

Institutions that wish to use cloud computing must work with their cloud service provider to devise an appropriate security plan that meets the general dbGaP Information Security Best Practices as well as these additional requirements that derive from the nature of multi-tenant clouds with default access to the internet. Please refer to the specific cloud service provider for methods, processes and procedures for working with controlled-access data subject to the GDS Policy in the cloud.

General Cloud Computing Guidelines

- Use end-to-end encryption for network traffic. For example, use Hypertext Transfer Protocol (HTTPS) sessions between your device and your data. Ensure that your service uses only valid and up-to-date encryption certificates.

- Encrypt data at rest with a user's (not cloud provider controlled keys) own keys. The National Center for Biotechnology Information SRA-toolkit includes this feature; other software providers offer tools to meet this requirement.

- Use security groups and firewalls to control inbound traffic access to your instance. Ensure that your security profile is configured to allow access only to the minimum set of ports required to provide necessary functionality for your services and limit access to specific networks or hosts. In addition, allow administrative access only to the minimum set of ports and source IP address ranges necessary.

- Be aware of the top 10 vulnerabilities for web applications and build your applications accordingly. To learn more, visit Open Web Application Security Project (OWASP) - Top 10 Web Application Security Risks. When new Internet vulnerabilities are discovered, promptly update any web applications included in your virtual machine (VM) images as well as VM host operating systems. Utilize the ISPO provided Nexpose vulnerability scanner to validate proper security patch levels are maintained by using credentialed scans every 30 days.

- Review any locally controlled network devices using Access Control Lists (ACLs) to ensure direct Internet access to devices processing, transmitting, or storing NIH protected information is not allowed. Work with ITS-CORE to review and validate ACLs on ITS-CORE firewalls and network devices provide additional security controls to restrict direct access to computing devise processing, transmitting, or storing NIH protected information.

Audit and Accountability

- Ensure that data is accessible only to those approved for access, and controls for changing that access are retained by the investigator who submitted the DAR and the appropriate IT staff. A mechanism for monitoring and notification needs to be in place to monitor permission changes. The IT data custodian should have either manual (periodic visual review every week) or systems such as a log review service or Security information and Event Management (SIEM) appliance to automate this process.

- Ensure that account access is logged along with access controls and file access and this information is reviewed by the investigator on a regular basis to ensure continued secure access. Consult with ITS-ISPO for possible cloud based services to meet this control.

Computer Image Specific Security

- Ensure images do not contain any known vulnerabilities, malware, or viruses. A number of tools are available for scanning the software, such as Chkrootkit, rkhunter, OpenVAS and the FSU enterprise supported Systems Center Endpoint Protection (SCEP).

- Ensure that Linux-based Images lock/disable root login and allow only sudo access. Additionally, root password must not be null or blank.

- Ensure that images allow end-users with OS-level administration capabilities to allow for compliance requirements, vulnerability updates, and log file access. For Linux-based Images, this is normally through SSH, and for Windows-based virtual machine images, this is normally through RDP.

Best Practices for Specific Cloud Service Providers:

Examples of cloud service provider best practices are provided in the links below; links to the best practices of additional cloud service providers will be appended to this document when they become available. Please be aware that these are provided for convenience only, and do not imply endorsement by the NIH or the United States Government for any of these services, nor does the government guarantee that these links lead to the most current version of these best practices. NIH recommends that investigators consult with their cloud service provider to ensure that they are using the most up to date best practice documents.

Amazon Web Services:

- http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AMIs.html
- http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-best-practices.html
- http://aws.amazon.com/documentation/ec2/

Google Cloud Platforms:

- https://cloud.google.com/developers/articles/best-practices-for-configuring-permissions-on-gcp

Additional Resources for Testing and Best Practices

Center for Internet Security (CIS)

The Center for Internet Security (CIS) (http://www.cisecurity.org/) is the only distributor of consensus best practice standards for security configuration. The *Security Configuration Benchmarks* are widely accepted by U.S. government agencies for Federal Security Information Act (FISMA) compliance, and by auditors for compliance with the International Organization for Standardization (ISO) standard as well as the Gramm-Leach-Bliley (GLB) Act, Sarbanes-Oxley (SOX) Act, federal Health Insurance Portability and Accountability Act (HIPAA), Family Educational Rights and Privacy Act (FERPA) and other regulatory requirements for information security. End user organizations that build their configuration policies based on the consensus benchmarks cannot acquire them elsewhere.

National Institute of Standards and Technology (NIST)

NIST, an agency of the US Department of Commerce provides information security standards and best practices for the federal government. The NIST Special Publications (SP) and Federal Information Processing Standards (FIPS) provide useful and concrete guidance to users of information technology systems (http://csrc.nist.gov/publications/).

United States Government Configuration Baseline (USGCB)

The United States Government Configuration Baseline (USGCB) (http://usgcb.nist.gov) provides security configuration baselines for information technology products widely used across the federal government including desktop computers.

Appendix A

dbGaP System Security Plan (SSP) FAQ & Plan Template
System Security Plan Template

1. Information System Name/Title
[Enter the name of the system (or systems)]

2. Information System Owner
[Enter the name and contact information for the system owner]

3. Other Designated Contacts, Including Those with "root" Access.
[Enter the names and contact information for any other critical technical or administrative contacts for this system. This should include the IT (policy) director, system administrators, data center contacts, etc.]

4. Assignment of Security Responsibility
[Who is responsible for implementing security policy? Enter the name and contact information of the security contact for this system, if different from above]

5. General System Description/Purpose
[Please describe the system and its purpose. Is this a standalone system, a compute farm, shared use system, desktop PC? What is the operating system, version? What is the data storage capacity?]

6. Physical System Environment
[Where is this system maintained? Data Center, Lab? What physical access controls exist to secure the system? For example, is there a defined list of people with physical access to the system? Access record keeping system? Locking system? Alarm systems? Video surveillance?]

7. System/Network Diagram
[Insert a diagram of the relevant portions of the systems to convey system and network architecture. Please include relevant off-site links, data storage locations, user-access points and firewall locations.]

8. System Interconnections/Information Sharing [1]

8.1. Security Controls
[What security controls are in place to protect the data and system? What encryption will be used for data stored on portable devices or laptops? How are security patches and updates applied?]

8.2. Access Control
[How is access control implemented to restrict access to these data to those authorized and how are data protected from being copied to unapproved locations?

What protections are in place to identify, authenticate and control external user access?]

8.3. Awareness and Training
[What is the process to ensure all users have had the necessary computer systems security training and acknowledge the sensitivity of access to these data?]

8.4. Configuration Management
[As system configurations change, what tracking is in place to ensure that security is maintained?]

8.5. Auditing and Accountability
[How are IT infrastructure and security audited? How frequently? Are audit records maintained and protected?]

[1]For further guidance on minimum security controls, see NIST FIPS publications 199 (http://csrc.nist.gov/publications/fips/fips199/FIPS-PUB-199-final.pdf)