TEXAS A&M UNIVERSITY - TECHNOLOGY SERVICES

National Institutes of Health - Summarized Security Best Practices for Controlled-Access Data Checklist

□ Files on local infrastructure are **never** exposed to the internet.

□ Data posted on servers in any fashion are **never** made publicly available, e.g.*, investigator's (or institution's) website*.

□ Strong authentication technology is being used for access control. Two factor authentication technologies are preferred. When using a single factor password, set policies that mandate the following requirements:
  - Minimum 12 length characters
  - Does not contain usernames, 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
  - Password should expire every 120 days or at the rate required by institutional policies, whichever is more frequent.

□ If data must be placed on mobile devices, it is encrypted. Avoid allowing users to place controlled access data on mobile devices or removable media such as USB thumb drives (except where such media are used as backups and follow appropriate physical security controls).

□ All software patches are up to date.

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

□ Physical access to all servers, network hardware, storage arrays, firewalls and backup media are restricted to only those that are required for efficient operations.

□ Access to secure facilities is logged.
  - Ideally with electronic authentication.

□ Servers are kept from being accessible directly from the Internet (i.e., it must be behind a firewall or not connected to a larger network).

□ Unnecessary services on servers are being disabled.

□ The principle of Least Privilege is being enforced to ensure that individuals and/or processes grant only the rights and permissions to perform assigned tasks and functions, but not more.

□ Controlled-access genomic and phenotypic data on the systems are being secured from other users (directory permissions are restricted to only the owner and group). If exported via file sharing, limited access to remote systems is being ensured.

□ If systems are being accessed remotely, encrypted data access is being used (such as Secure Shell (SSH) or 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 being used on multiple systems (such as a compute cluster), then data access policies are being retained throughout the processing of the data on all the other systems.

□ If data is cached on local systems, directory protection is being kept and data is being removed when processing is complete.

□ Approved users are retaining the original version of the encrypted data, tracking all copies or extracts and ensuring that the information is not divulged to anyone except authorized staff members at the institution.

□ Collaborating investigators from other institutions have submitted an independent Data Access Request (DAR) and been approved by NIH to access the data. Outbound access from devices that host controlled access data has been restricted.

□ Data downloaded from NIH-designated data repositories is being destroyed is 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 or specified records within the context of media containing multiple records).

□ Only encrypted copies of the minimum data necessary are being retained at the institution to comply with institutional scientific data retention policy, as well as any data stored on temporary backup media as are required to maintain the integrity of the institution's data protection program.
  - If retaining the data on separate backup media is not possible, 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 is being deleted at the appropriate time, according to institutional policies.

□ Hard copies, CD ROMs, and other non-reusable physical media are being shredded.

□ Electronic files are being deleted securely.
  - ○ For personal computers, the minimum would involve deleting files and emptying the recycle bin or equivalent with equivalent procedures for servers.
  - ○ Optimally, use a secure method that performs a delete and overwrite of the physical media that was used to store the files.

□ Backups are reused (data deleted) and any archive copies are also being destroyed.