Data Retention Policy

Document Name: NCSA Data Retention Policy
Version: 1.0.1
Accountable: James Eyrich
Authors: Adam Slagell

Reviewed: October 16, 2023
Approved: March 7, 2016

Scope

This data retention policy covers all shared systems providing services to NCSA staff or customers. Currently, these are the systems operated by the Integrated Cyber-Infrastructure directorate at the NCSA. More specific system data retention policies and contracts/agreements can supplant this default policy.

Data Types

There are six major types of data considered: Staff, User, Project, Log, Temporary, & Business data.

Staff Data

Staff data are information on these shared systems, which people are given access to by virtue of being a staff member at the NCSA. Examples include AFS home directories, laptop/workstation backups, user space on the file server, personal spaces on the wiki, and list archives for lists that staff person owns.

User Data

User data are information on these shared systems, which people are given access to by virtue of having an allocation on an NCSA operated system. Examples include home directories on clusters, personal directories on storage services, and other storage that is given to users who are not necessarily sponsored guests or staff.

Project Data

Project data are information on these shared systems, which are shared by a project and should persist beyond any one person leaving that project. Examples include web server spaces in AFS, project spaces on clusters or archives, group or project folders on the file server, project wiki spaces, and project email list archives.

Logs & Accounting Data

Logs and accounting data are metadata generated by these shared systems to assist in the maintenance, operation, and accounting of those systems. These include system and network logs, allocation usage, security alerts, failure notices, etc.

Temporary Data

Temporary data can be generated or created by users in special folders for the express purpose of short term use. Examples include scratch spaces on clusters, temporary and swap folders used by operating systems, and data staging areas used to transfer data to another system or person.

Business Data

Business data are financial, contract, personnel or other data required by the administrative staff at NCSA to do their daily job. This data may rest on centrally managed workstations or laptops, file servers, wikis, or backups. Business data sitting on non-NCSA University systems are not considered here.

Retention Policies

These timelines are not a promise to cleanup data that may be older, but only a commitment to how long we **minimally** retain such data. It is important to note data life span can be cut short by decommissioning or failure of the system hosting it. It is not a promise to backup any data elsewhere, either.

Staff Data

Staff data is retained for 180 days after employment ends, whereas, their user data on allocated systems will remain as long as they have an active allocation on the relevant systems.

User Data

User data remains 90 days after the user has no allocation on the relevant systems. The user may no longer have access to the system, in which case NCSA staff can help retrieve it until expiration.

Project Data

Project data is retained for one year after the end of a project, typically the end of its funding or retirement of its resources.

Logs & Accounting Data

System and network logs are retained for one year, unless they are a part of an incident investigation in which case they are held indefinitely. Accounting and user profile data in the IRG databases are never deleted.

Temporary Data

Temporary data are only kept for 48 hours by default, and may be cleaned up automatically after that time.

Business Data

Business data are kept as long as required by law or University policy, and this differs for each kind of contract, personal and financial data. NCSA system operators with the proper administrative unit(s) to support their requirements for data retention.

Privacy

As a unit of the University of Illinois at Urbana-Champaign, we must abide by the University's policies regarding privacy, which includes the Web Privacy Notice on UIUC web pages.

Any data may be accessed in the course of a security incident or to provide for the operation of services, e.g., troubleshooting system issues. However, efforts are made to notify users or staff if their home directories are accessed in the process of providing support for our daily operations.

Staff should have no expectation or privacy when storing data on NCSA systems, and managers along with HR can decide when it is necessary to investigate the contents of staff data. ~~However, staff must be notified after the fact.~~

Unless there is an exception for incident investigation or system operation, user data remains private and can only be shared with permission of the data owner.

Project data is owned by the Principal Investigator or project lead, and they decide who has access to the data.

Log data is not considered private, unless there is special policy regarding the specific type of data. NCSA may share log data with researchers that have IRB approval to utilize the data. Similarly, temporary data is not considered private by default and is often stored in shared spaces.

Finally, business data is shared only with those who need it to do University business, and it may be protected by additional contracts, laws, and regulations.

Exceptions

This policy must be overridden if it is in conflict with any law or University policy. More specific policies for particular systems and agreements/contracts with third parties will also supplant this default retention periods of this policy.

Exception to the privacy provisions, however, must be approved by the Internal Infrastructure Board.