WELLESLEY

Central File Storage Policy for Faculty/Staff

1.0   Policy Statement

The *Wellesley College Central File Storage Policy for Faculty/Staff* outlines the multiple central file storage options available for Wellesley College faculty and staff in order to address the growing need for greater file storage.

2.0   Overview

As the use of technology for scholarly purposes and the need for collection of data continues to increase, the need for greater file storage has increased as well. For example, some faculty have very large quantities of data (e.g., research sets) that are too large to store on a local disk or central file server. When a centrally managed file storage option is not sufficient, some Wellesley College faculty and staff have opted to manage their own file server.

3.0   Purpose

Distributed file server management outside of Library & Technology Services (LTS) creates multiple potential issues: insecure storage of data; incompatibilities with approved computer configurations; little or no back-up of data; duplication of stored data; inability to share data among departments; and issues with providing hardware support. This policy intends to address these concerns through offering centrally managed file storage options that meet the needs for both small and large data storage.

4.0   Scope

This policy applies to all Wellesley College faculty and staff.

5.0   Policy
5.1   Responsibilities

Library and Technology Services will provide secure file storage space, maintain the servers and provide back-up for the data. The Chief Information Officer has responsibility for this policy.

5.2   General Policy

Information Services provides the following options for faculty and staff:
- On campus storage on a file server - NTM
- Sakai Resources

- On campus large file storage
- Off campus large file storage on Amazon S3

On Campus storage

Each Wellesley College faculty and staff member will be eligible for 7 GB of storage on NTM. Additionally, each department will be eligible for 30 GB of shared storage on NTM. NTM storage is considered "primary storage" in that it is located on high performance, and highly redundant, disks. As this space is limited, Wellesley College will strictly enforce quotas stated above. The individual employee space is to be used for important documents that can be accessed from multiple locations and the department space is to be used for sharing and collaborating with others at Wellesley College. Through SSL VPN, NTM is accessible from home and other remote locations, which makes this is a viable shared storage option for many employees.

Sakai Resources

Faculty members can use Sakai Resources to store documents used for teaching, and all employees can use MyWorkspace and Project Sites for personal use and collaboration respectively. Each of these Sakai sites has a default storage quota of 1 GB.

On Campus Large File Storage

Several faculty members engaged in research have a need to store very large (Terabytes) quantities of data. In some cases, these data sets are actively used, and in other cases, they are archived. NSF now requires that the College have a Data Management Plan that needs to be specified in the grant proposal. This plan would outline how the College will disseminate data related to research to the public, if funded.

LTS is developing a practical solution to this requirement that involves using "secondary storage" and a chargeback model. Secondary storage refers to disks that are not as fast as primary storage, and are not as redundant. LTS will offer to back up the most recent copy of the data. The reason for the chargeback is to discourage misuse of the data storage option. The monies collected through the chargeback will be used to add more storage space as needed. LTS is in the process of restructuring available storage for this purpose.

The recommended charge for secondary storage is $500 per terabyte of storage for three years. This is a reasonable cost compared to marketplace value, along with the added benefit that the data is backed up and stored in an environmentally controlled data center.

Off Campus File Storage in Amazon S3

An alternative to on-campus file storage is moving that data to "the cloud." The cloud refers to a series of connected servers distributed throughout the country that are accessed through the internet. There are many instances where the requirements for large data file storage can be easily met by off site storage in the cloud. LTS recommends the use of Amazon S3 because it is one of the most reliable services. There are considerable unknowns about this option with respect to the transfer speeds of large data sets, the cost (which depends on the data transfer), convenience for access, etc. However, overall it will be a less expensive option because it is not necessary to purchase any hardware or allocate

staff resources for maintaining it. LTS will administer the program for anyone who wishes to use it. There will be a full chargeback to the faculty or department for use of Amazon S3.

LTS has also commissioned a test system on Amazon EC2. For those faculty wishing to run data analysis on data stored in the cloud, LTS can potentially commission another computer in the Amazon Cloud to run the software (assuming that the software license allows for this). If the data contains human subject data or labor statistics, this model will not be viable due to security concerns.

For the storage of research sets that involve chargeback, LTS will work with the faculty to determine if the cost of storage can be built into the grant proposals,

6.0   Enforcement

Misuse of the data storage options listed above or any other failure to comply with this policy could result in loss of privilege to use the space. Storage space options are subject to change as the technology that supports them changes and newer technologies become available. Wellesley College reserves the right to change the amount of chargeback for storage space at any time it deems necessary.

7.0   Effective Date

Approved 10/26/2011 and effective immediately.