

WHAT FACULTY SHOULD DO DURING DATA CLEANUP

During the Data Cleanup campaign faculty are encouraged to locate and then securely delete or archive any sensitive data files found on their computer. If time permits, sensitive documents should also be located and depending on the retention requirement for the document it should be securely destroyed, or archived.

To determine how long a sensitive and non-sensitive business, financial, student, or research related document or file must be retained by the Institute, faculty can reference the Board of Regents (BOR) data retention guidelines, which are located on the Resources page of the data cleanup website.

It is important that faculty is aware of the sensitivity level and value of the Institute data that may be stored on their computer and in the office. Faculty can determine the classification of data, by reading and becoming familiar with the Georgia Tech Data Security Classification Guide. The guide is designed to be used as a reference tool and to provide tips and other general information regarding specific data types. A complete copy of the guide can be found on the Resources page of the Data Clean Up website www.datacleanup.gatech.edu.

Faculty are encouraged to work closely with their CSR to locate the sensitive files on their computer, then determine if the files should remain on the computer or be securely deleted or archived to the department's or another specified fileserver for faculty and/or research data. If one is not available, consider identifying or purchasing one, or archive the sensitive files that are not used frequently to a CD/DVD or other removable media and secure it in a safe location.

Sensitive files containing social security, credit card, and gtID numbers can be located on the computer by running a discovery tool name Cornell Spider. The tool was created by Cornell University, but modified by Georgia Tech to work more efficiently in our environment. Your CSR can either install the tool by visiting your computer or by using a centrally managed console such as LanDesk.

Important Notes:

1. Other sensitive files that a faculty member may have are grade books/spreadsheets, grant forms, and research documents.
2. A file containing a list of student's name along with their letter grade is considered sensitive information. Since Spider is designed to locate specific strings of data, such as social security, credit card and gtID numbers the tool **will not** locate this combination of data.
 - a. Such files can be found by selecting the ***“Documents (word processing, spreadsheet, etc.)”*** option using the Microsoft Search Companion.
 - i. Enter ****.xls*** to find all Excel files, or
 - ii. Enter ****.doc*** to find all Word files

Note: This will list all Excel or Word files on the computer. If you have an idea of where this combination of data may reside on your computer, you can narrow the search by selecting the ***“Use advance search options”*** to search a specific directory or folder.

Once the Spider tool has been installed and ran on the computer, the faculty member should do the following:

1. Review the report and follow the path of the file listed in the report to locate each file identified as containing sensitive data.

Note: Faculty may have to work with their CSR to assist with following the path to locate the sensitive files.

- a. Once located faculty will have to determine if they need to securely delete, archive, or keep the files on their computer.
 - i. The files should be securely deleted if old and/or not needed for current courses, research projects, assignments etc.
 - ii. The files should be archived to a fileserver if one is available and the files are needed.
 - iii. The files should be archived to removable media such as a CD/DVD, encrypted USB drive etc. and stored in a securely controlled location such as a locked desk drawer, a safe, or file cabinet, if a fileserver is not available however and the files must be retained for the current job function.

Note: The files should only remain on the computer if they can not be secured on the fileserver or removable media, or if the Dean approves keeping the sensitive data on the computer. Systems that must continue storing sensitive data should be

encrypted using the GT approved PGP whole-disk encryption application. A license for the application will be issued by OIT to all computers approved by the Dean to continue storing sensitive data.

- b. As faculty locate and review their sensitive files, they should also update the Spider report to indicate how they handled each file (i.e. securely delete, archive, or keep).

The Spider report should be pulled into an application such as Microsoft Excel by the CSR or faculty member, so that additional information can be added to the report by the faculty member to indicate whether or not they securely deleted, archived, or need to keep the sensitive file.

- c. Once the faculty member has viewed each file and updated the report, the report should be reviewed by the Dean for approval of the sensitive information distribution (*i.e. securely delete, archive, or keep*).

Important Notes:

1. The CSR or faculty member may choose to verify that sensitive files that are not needed have been removed from the computer by running Spider again and comparing the reports.
2. The spider log file contains information to the location of files that contain sensitive data. It is important that the CSR move the file off the system as soon as possible.

FOR MORE INFORMATION REGARDING DATA CLEANUP, VISIT www.datacleanup.gatech.edu
FOR SUPPORT OR QUESTIONS, E-MAIL datacleanup@gatech.edu