

Spider 2008

Spider 2008 is an automated tool developed by Cornell University to identify sensitive data that may be located on a computer. When Spider runs, it completely scans all files and directories that are specified by a user and looks for sensitive data by using simple pattern matching (e.g., XXX-XX-XXX = SSN). Though this method makes Spider very thorough, false positives are very common.

Prior to using Spider, it is important to be aware of the following:

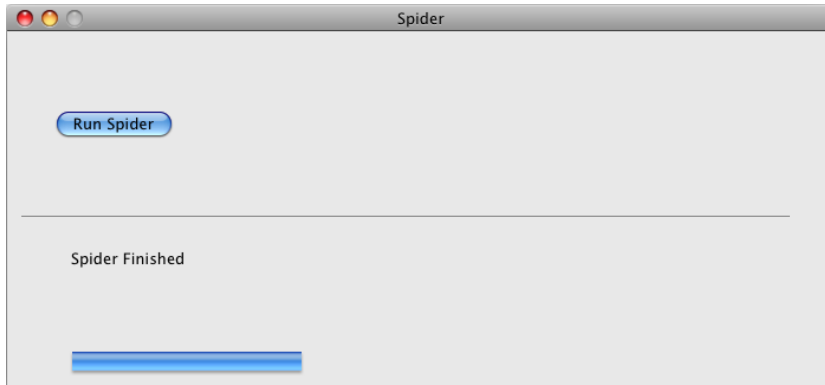
- **False positives** - Spider is doing simple pattern matching (e.g., XXX-XX-XXXX = SSN), which means false positives are very common. This means that a number using the same format XXX-XX-XXXX will be flagged by Spider as an SSN. You will need to review and interpret the results critically.
- **False negatives** -If SSN or credit card data is entered in a non-standard format, Spider will not identify it. For example, "SSN:123456789" is not found with Spider's default configuration. Along similar lines, remember that Spider is only looking for SSN & credit card data. You may have other types of sensitive data on the machine.

Running Spider

- 1) Find the Spider program on your Dock or Applications folder and open the program.
- 2) Start scanning your computer by pressing the “Run Spider” button.



3) Once Spider has finished scanning the main screen will display that “Spider Finished”.



4) After Spider has finished, it will place a log of files with suspected sensitive data on your desktop named “spider.log”



5) When you open the log it will display a list of files that Spider suspected may contain sensitive data. It is important to review this list because Spider will report many false positives.

