

Bates Numbering - *What's in a number anyway?*

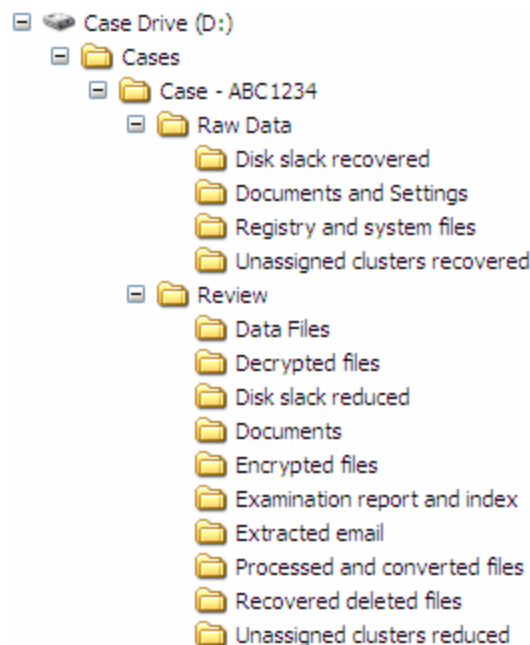
Abstract

Most experts agree the computer forensics process closely follows the phases of discovery in support of civil litigation. The phases of civil discovery are Identification, Preservation, Filtering and Production. This paper focuses on the Production phase of computer forensics in support of civil litigation.

Implementation

During the Production phase, the forensics examiner is faced with the challenge of producing a wealth of data in an organized manner to the legal team. The data produced must be easily referenced by all persons concerned.

By taking a look at the directory structure created by a computer forensics examiner in a hypothetical case we can see early stages of organization. In this case the forensics examiner created a case work directory on disk containing the following sub-directories:



By creating the above directory structure early, the forensics examiner is not only supporting analysis, they are providing groundwork for the final “Production” phase of the case.

Note: Raw data sub-directories will vary based on case, examiner preference and tools used. Many of today’s integrated computer forensics tools negate the need for extensive raw data directories.

Today’s desktop computer can contain as many as 50,000 or more files. With such a large number of files, printing out even the reduced evidence files can become unmanageable. Most forensics examiners will prepare a three-ring binder with a printed version of the case report and possibly a sampling of the most interesting evidence then copy the remaining review directories to CD-ROM.

If you will recall earlier it was suggested “The data produced must be easily referenced by all persons concerned.” To aid in referencing of evidence, the legal profession has long used a system called “Bates numbering” to refer to documents. Bates numbering is simply a way to refer to evidentiary documents by number rather than name. Not only is Bates numbering helpful when referring to a large number of documents, with electronic documents Bates numbering eliminates issues related to two documents with the same name. The Bates number is normally a series of letters followed by sequential numbers.

Note: The term “Bates Number” comes from the Bates Manufacturing Co. which was incorporated on September 13, 1890 in New York State. The Bates Manufacturing Co. manufactured and sold automatic hand-held numbering machines. The company was bought by Edison Phonograph Works (*Thomas Edison’s company*) in 1892.

In computer forensics many examiners implement bates numbering prior to copying evidence files to CD-ROM. This is achieved by renaming files with the Bates number added to the original file name as shown in the example below:

Pre-bates file name = MyFile.txt
Post-bates file name = MyFile.AAAANNNNN.txt

Where AAAA is a series of unique characters and NNNNN is a series of sequential numbers. The file above might be renamed to "MyFile.TP00001.txt".

In cases where the original file has no extension, three underscores would be added, thus the file "MyFile" would become "MyFileTP00001.____"

Note: Most integrated computer forensics tools provide the ability to create Bates numbered files. ProDiscover DFT supports the Bates numbering method described above, which was originated by Troy Larson.

In large civil cases, evidence files may span many CD-ROMs. In these cases bates numbering can become more involved by creating a directory structure that becomes part of the Bates numbering. The naming method described below is designed for such large cases.

CD File and Directory Format

The following is an example of a full path and file name using this method:

TP180000\00\01\02\94.doc

TP180000 is both the CD-ROM volume name and the main directory for the image file.

Each image file path in the example above will contain three subdirectories each in turn containing 100 subdirectories, named from 00 through 99. The number of document files in each path will be 100 except for the first directory, which will start at 01 and go through 99, and the last directory, which may contain fewer than 100 document files.

The file name in the above example is "94.doc" wherein ".doc" is the file name extension. The file extension could just as easily have been ".DBX" or ".JPG".

Cross Reference File format

A cross-reference file would be created for each CD-ROM and delivered on diskette and CD-ROM. A cross-reference file is an ASCII text file containing the Bates number of each document on the CD-ROM with its corresponding document address.

The name of the file should be the CD-ROM volume name followed by ".CRF"

Example: TP180000.CRF

Assume that a given document is recorded on CD-ROM Volume TP180000. The corresponding document's cross-reference Bates number file entry might look like:

ABCDE001-TP180000\00\00\00\01.doc

Note: The method described above does not support evidence where it's location on disk is important to the case. For such cases a modification to the format would need to occur. The original file location on disk could be added to the cross-reference file.

For large cases it is common practice for law firms to establish Bates numbering guidelines that fit the scope and type of evidence expected. Always check with the contact attorney for each case to see if they have pre-established Bates numbering formats for your case.

Conclusion

The Bates numbering created during the Production phase is an integral part of providing reference to evidence for civil litigation. While there is no regulation requiring a specific Bates numbering format be used, the formats mentioned in this white paper will help in constructing a method of clearly referencing evidence.