

CyberBytes



DATA FILES — Legal Definitions

Active Data

“Active data” is the information currently accessible on a computer, such as word processing documents, spreadsheets, databases, e-mail messages and electronic calendars. Generally, active data is relatively simple to access through the use of a computer’s file manager program. It can be found on an individual’s office desktop computer, laptop, home computer, an assistant’s computer, a PDA and a network file server. Moreover, because users frequently create special files or folders in which to store e-mails or other electronic documents that pertain to a particular subject matter, active data will usually be fairly easy to sort for relevant information. Most computer programs also contain search engines that can be used to narrow the scope of potentially relevant documents.

Replicant Data

“Replicant data” (or “archival data”) is the information a computer automatically backs up as you work on a file. These backed up files are created and saved in order to recover data that may be lost due to a malfunction or power loss. Replicant data is useful because it creates a copy or several copies of a document that the user may not erase. In fact, the user may not even be aware of these “file clones” because they are generally stored in a different directory than active data. On most networked systems, this Replicant data is stored on the hard drive as opposed to a centralized network file server. Consequently, a document, or part of it that was purged from a server, may be retrievable from a user’s hard drive.

Backup Data

“Backup data” is information copied to a removable medium in the event of a system failure. Most businesses have their networks backed up on a routine schedule, while individual users may or may not backup their information. Thus, one can find backup data on system-wide backup tapes, recovery backup tapes that may be stored off site, and on personal backups such as computer disks.

Backup data is particularly useful in that it provides historical snapshots of the data stored on a system on the specific day that the backup was made, allowing one to obtain information regarding the progress of a matter. On the flip side, because backup tapes contain large amounts of data, it is frequently time consuming and expensive to restore this data to review the material pertinent to your case.

Residual Data/Inactive Data

Unlike general “paper” discovery, electronic documents thought to be lost or destroyed are more often than not recoverable; yielding what is often an untapped source of information in a case. Simply pressing the “delete” button does not mean that the document is no longer in the computer. “Residual data” is information that is actually recoverable even though an attempt has been made to “delete” the document. When a file is “deleted” the computer makes the space occupied by that file available for new data. Unless that space is “re-written,” the so-called deleted document is generally recoverable by using “undelete” or “restore” commands contained in some systems’ operating software or through other programs”

[Residual Data: Residual Data (sometimes referred to as (“Ambient Data”) refers to data that is not active on a computer system. Residual data includes (1) data found on media free space; (2) data found in the file slack space; and (3) data within files that have functionally been deleted in that it is not visible using the application with which the file was created, without use of the undelete or special data recovery techniques. The Sedona Principles (2004)

In addition, the following classifications of electronic information may be important in your cases.

- Ambient data – generally referred to as “residual data.” See description above
- Archival and legacy data – usually stored on backup tapes in formats that may or may not be easy to access.
- Forensic, bitstream or clone copy – is an exact bit-by-bit copy of a hard drive of a computer system. Note: the use of the term “mirror copy” is frequently used by attorneys in lieu of bitstream copy in their discovery requests. This is a complete error on their part, because most responding party IT managers assume that the mirror copy is meant to be only a copy of all active files on a hard drive which will exclude all inactive or deleted files possibly being expected in production.
- System data – “System data, or information generated and maintained by the computer itself. The computer records a variety of routine transactions and functions, including password access requests, the creation or deletion of files and directories, mainframe functions, and access to and from other computers, printers, or communications devices.”