

Glossary

Active Data	Files that can be seen on a computer by the average user. Does not include deleted files or system files that may be hidden.
ASCII	American Standard Code for Information Interchange. Standard text with no formatting which means it can be read by most computer programs.
Backup tapes	Backup tapes are a form of media used to store data. Some of the most common types of backup tapes are DAT, DLT and LTO. All tapes are NOT created equal. A DAT tape usually holds a few hundred MEGABYTES whereas an LTO tape can hold a few hundred GIGABYTES (1024 MBs = 1 GB).
Bit-Stream Copy	Also called an evidentiary image. This is an exact sector by sector copy of a hard drive which allows for retrieval of deleted files.
Byte Kilobyte Megabyte Gigabyte Terabyte	Unit of measurement for digital data. One byte = one character. 1024 bytes = 1 kilobyte (KB) – This document is 36 KB. 1024 KB = 1 megabyte (MB) – A floppy disk holds 1.4 MB. 1024 MB = 1 gigabyte (GB) – About 731 floppy disks. 1024 GB = 1 terabyte (TB). About 213 standard DVDs.
Computer Forensics	Recovering deleted files and determining the activity of a computer user.
De-Dupe	Shorthand for de-duplication. The concept of removing all duplicate documents seems straight-forward, but actually raises many questions. Is the same email in two different inboxes truly a duplicate? And what process is used to de-duplicate? The standard method is by using hash values (see below) but this is generally not very effective on email.
Electronic Discovery	Acquiring digital media for production in litigation. Has also been used to describe the process of converting paper files to digital media (e.g., TIFF and OCR files).
Evidentiary Image	Same as a bit-stream copy. An exact sector by sector copy of a hard drive that allows for retrieval of deleted files.
File extension	Typically the last three characters of a file name after the period that indicates the program or application used to create the document. For example, Word documents end with a .doc file extension, Excel spreadsheets have a .xls file extension and Word Perfect uses .wpd file extension.
Hash value	The “DNA” or fingerprint for a digital file. A string of characters that is unique to that particular file or group of files. MD-5 and SHA are the two most common algorithms used to create hash values. Hash values can be used to de-dupe files or to confirm authenticity.
Image	This word can have numerous meanings in the electronic discovery and computer forensics world. An image is a picture and is often used to describe TIFF files created from paper documents. But it is also used to describe a copy of a hard drive or other digital media (such as a Ghost image).
Lotus Notes	Made by IBM. One of the two most common corporate email programs. The other is Microsoft’s Outlook.
Metadata	Data about data. Information such as the author, date created and date

	modified. There is file or system metadata and document or application metadata. File metadata is assigned by the computer's operating system whereas document metadata is assigned by the application or program used to create the file. If a file is moved from one location to another the file metadata will change, but most of the document metadata will remain intact.
.msg file	An individual email file, commonly associated with Outlook or Outlook Express.
Native File Format	A file that is in its original format (e.g., Word document, Excel spreadsheet, .pst file, etc.).
.nsf file	A database store for Lotus Notes. Typically contains email, contacts, calendar items and journal entries. Similar to Outlook's .pst file
OCR	Optical Character Recognition. A software program converts the text of a printed document to digital format so the document can be searched. Accuracy rates can vary widely depending on numerous factors such as the quality of the original and the program used for scanning which obviously impacts the validity of any search results.
Outlook	One of the two most common corporate email programs (the other is Lotus Notes). Made by Microsoft and needs to be purchased, but typically comes as part of one of the Microsoft Office packages.
Outlook Express	An email program that comes standard on all Microsoft operating systems. Depending on their email provider, a user could download their web-based email and view it with this program.
.ost file	Offline Storage Table. Similar to a .pst file but typically used by remote users that will read and respond to email offline and then synchronize it with their company's Exchange server when Internet access is available.
.pst file	Personal Stores Table. The compressed file used by Outlook to store an individual user's email, contacts, calendar items, journal entries and notes. Similar to Lotus Notes' .nsf file.
Slack Space	A portion of the hard drive between where one file ends and another file begins. This can be an important place to look if searching for old deleted data.
TIFF or TIF	Tagged Image File Format. A graphic file that is commonly used by litigation support vendors to create pictures of either digital or paper documents.
Unallocated Space	The portion of the hard drive that is not allocated to active files. When a file is deleted, only the link to the file is removed. The actual file remains in the unallocated space of the hard drive until it has been overwritten.
Web-based email	As the name implies, email that is accessed and stored on the Internet. This includes programs such as Hotmail, Gmail, Yahoo and others. Typically these programs store the email on their own servers, but sometimes the email can be moved down to the user's computer.