
E-Discovery Rule Amendments: Have they created the need for you to learn a new language?

Most students in the United States start learning a foreign language in high school or late middle school. Most of us who consider our formal education to be a distant memory - especially our foreign language requirements - will be surprised (and maybe irked) to learn that the federal and state rule amendments addressing electronically-stored information have created a new need for judges and lawyers and their clients to learn a new "language."

It is a fact of modern life that an enormous volume of information is created, exchanged, and stored electronically. Technology is redefining the legal profession and state and federal judiciaries have responded to this fact of life by specifically addressing these changes in their respective court rules of civil procedure. With these changes comes exposure to a whole new vernacular to the legal profession.

The purpose of this article is to give you a basic understanding of a small sampling of the technological terms you will likely confront as you face e-discovery issues. The terms and definitions are taken from The Sedona Conference Glossary: E-Discovery & Digital Information Management, A Project of The Sedona Conference Working Group on Electronic Document Retention & Production. While the Sedona Conference Glossary identifies and defines well over 600 technical terms, we endeavor to select and define thirteen of the most commonly used terms in the context of e-discovery. Here they are:

1. Data (electronic): Any information stored on a computer. All software is divided into two general categories: data and programs. Programs are collections of instructions for manipulating data. In database management systems, data files are the files that store the database information. Other files, such as index files and data dictionaries, store administrative information, known as metadata.

(a) Active Data: Information residing on the direct access storage media (disc drives or servers) of computer systems, which is readily visible to the operating system and/or application software with which it was created and immediately accessible to users without restoration or reconstruction.

(b) Backup Data (disaster recovery data): An exact copy of system data which serves as a source for recovery in the event of a system problem or disaster. Backup Data is generally stored separately from Active Data on portable media. Backup Data is distinct from Archival Data in that Backup Data may be a copy of Active Data, but the more meaningful difference is the method and structure of storage which impact its suitability for certain purposes.

(c) 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 file slack space; and (3) data within files that has functionally been deleted in that it is not visible using the application with which the file was created, without use of undelete or special data recovery techniques. May contain copies of deleted files, Internet files and file fragments.

For the complete Glossary, visit <http://www.thesedonaconference.org>. We encourage you to refer to the Glossary to the extent you confront terms not identified and defined herein.

2. **Deletion:** The process whereby data is removed from active files and other data storage structures on computers and rendered inaccessible except through the use of special data recovery tools designed to recover deleted data. Deletion occurs on several levels in modern computer systems: (a) file level deletion renders the file inaccessible to the operating system and normal application programs and marks the storage space occupied by the file's directory entry and contents as free and available to re-use for data storage; (b) record level deletion occurs when a record is rendered inaccessible to a database management system (DBMS) (usually marking the record storage space as available for re-use by the DBMS, although in some cases the space is never reused until the database is compacted) and is also characteristic of many email systems; and (c) byte level deletion occurs when text or other information is deleted from the file content (such as the deletion of text from a word processing file); such deletion may render the deleted data inaccessible to the application intended to be used in processing the file, but may not actually remove the data from the file's content until a process such as compaction or rewriting of the file causes the deleted data to be overwritten.

3. **File Format:** The internal structure of a file, which defines the way it is stored and used. Specific applications may define unique formats for their data (e.g., "MS Word document file format"). Many files may only be viewed or printed using their originating application or an application designed to work with compatible formats. There are several common email formats, such as Outlook and Lotus Notes. Computer storage systems commonly identify files by a naming convention that denotes the format (and therefore the probable originating application). For example, "DOC" for Microsoft Word document files; "XLS" for Microsoft Excel spreadsheet files; "TXT" for text files; "HTM" for Hypertext Markup Language (HTML) files such as web pages; "PPT" for Microsoft Powerpoint files; "TIF" for tiff images; "PDF" for Adobe images; etc. Users may choose alternate naming conventions, but this will likely affect how the files are treated by applications.

4. **Metadata:** Information about a particular data set or document which describes how, when and by whom it was collected, created, accessed, modified and how it is formatted. Can be altered intentionally or inadvertently. Can be extracted when native files are converted to image. Some metadata, such as file dates and sizes, can easily be seen by users; other metadata can be hidden or embedded and unavailable to computer users who are not technically adept. Metadata is generally not reproduced in full form when a document is printed.

5. **Native Format:** Electronic documents have an associated file structure defined by the original creating application. This file structure is referred to as the "native format" of the document. Because viewing or searching documents in the native format may require the original application (for example, viewing a Microsoft Word document may require the Microsoft Word application), documents are often converted to a vendor-neutral format as part of the record acquisition or archive process. "Static" formats (often called "imaged formats"), such as TIFF or PDF, are designed to retain an image of the document as it would look viewed in the original creating application but do not allow metadata to be viewed or the document information to be manipulated.

6. **Network:** A group of two or more computers and other devices connected together ("networked") for the exchange and sharing of data and resources. A local-area network (LAN) refers to connected computers and devices geographically close together (i.e., in the same building). A wide-area network (WAN) refers generally to a network of PC's or other devices, remote to each other, connected by telecommunications lines. Typically, a WAN may connect two or more LAN's together.

7. **RAM (Random Access Memory):** Hardware inside a computer that retains memory on a short-term basis and stores information while the computer is in use. It is the "working memory" of the computer into which the operating system, startup applications and drivers are loaded when a computer is turned on, or where a program subsequently started up is loaded, and where thereafter, these applications are executed. RAM can be read or written in any section with one instruction sequence. It helps to have more of this "working space" installed when running advanced operating systems and applications. RAM content is erased each time a computer is turned off.

8. **ROM (Read Only Memory):** Random memory which can be read but not written or changed. Also, hardware, usually a chip, within a computer containing programming necessary for starting up the computer, and essential system programs that neither the user nor the computer can alter or erase. Information in the computer's ROM is permanently maintained even when the computer is turned off.

9. **Server:** Any central computer on a network that contains data or applications shared by multiple users of the network on their client PC's. A computer that provides information to client machines. For example, there are web servers that send out web pages, mail servers that deliver email, list servers that administer mailing lists, FTP servers that hold FTP sites and deliver files to users who request them, and name servers that provide information about Internet host names.

10. **Slack/Slack Space:** The unused space on a cluster that exists when the logical file space is less than the physical file space. Also known as file slack. A form of residual data, the amount of on-disc file space from the end of the logical record information to the end of the physical disc record. Slack space can contain information soft-deleted from the record, information from prior records stored at the same physical location as current records, metadata fragments, and other information useful for forensic analysis of computer systems.

11. **System:** A system is: (1) a collection of people, machines, and methods organized to perform specific functions; (2) an integrated whole composed of diverse, interacting, specialized structures and sub-functions; and/or (3) a group of sub-systems united by some interaction or interdependence, performing many duties, but functioning as a single unit.

12. **Temporary File:** Temporary (or "temp") files are files stored on a computer for temporary use only, and are often created by Internet browsers. These temp files store information about Web sites that a user has visited, and allow for more rapid display of the Web page when the user revisits the site. Forensic techniques can be used to track the history of a computer's Internet usage through the examination of these temporary files. Temp files are also created by common office applications, such as word processor and spreadsheet applications.

13. **Unallocated Space:** The area of computer media, such as a hard drive, that does not contain normally accessible data. Unallocated space is usually the result of a file being deleted. When a file is deleted, it is not actually erased, but is simply no longer accessible through normal means. The space that it occupied becomes unallocated space, i.e., space on the drive that can be reused to store new information. Until portions of the unallocated space are used for new data storage, in most instances, the old data remains and can be retrieved using forensic techniques.

James J. Shrager, Esq., Margaret Raymond-Flood, Esq. and Fernando M. Pinguelo, Esq. are members of Norris McLaughlin & Marcus and head the firm's Response to Electronic Discovery and Information (REDI) Group. In addition to focusing their practice on complex business litigation matters, they regularly advise clients and publish and speak on issues related to electronically stored information, electronic discovery, and records management practices. If you have any questions or concerns about electronically stored information, or any other questions, please contact them at e-discovery@nmmlaw.com or 908-722-0700.

The *E-Discovery Law Alert* provides information to our clients and friends about current legal developments or general interest in the area of e-discovery law. The information contained in this *Alert* should not be construed as legal advice, and readers should not act upon such without professional counsel. Copyright © 2007 Norris McLaughlin & Marcus, P.A.