

Exhibit A-2

Intranet: A private network that uses Internet-related technologies to provide services within an organization.

IP address (Internet Protocol address): A string of four numbers separated by periods used to represent a computer on the Internet - a unique identifier for the physical location of the server containing the data. See TCP/IP (e.g., 206-1432.001).

IPX/SPX: Communications protocol used by Novell networks.

IS/IT Information Systems or Information Technology: Usually refers to the people who make computers and computer systems run.

ISA: Industry Standard Architecture.

ISDN (Integrated Services Digital Network): An all digital network which can carry data, video and voice.

ISIS and TWAIN Scanner Drivers: Specialized applications used for communication between scanners and computers.

ISO 9660 CD Format: The International Standards Organization format for creating CD-ROMs that can be read worldwide.

ISO: International Standards Organization.

ISP (Internet Service Provider): A business that provides access to the Internet, usually for a monthly fee. ISPs may be a source of evidence through files (such as ISP email) stored on ISP servers.

IT (Information Technology) Infrastructure: The overall makeup of business-wide technology operations, including mainframe operations, standalone systems, email, networks (WAN and LAN), internet access, customer databases, enterprise systems, application support, regardless of whether managed, utilized or provided locally, regionally, globally, etc., or whether performed or located internally or by outside providers (outsourced to vendors). The IT Infrastructure also includes applicable standard practices and procedures, such as backup procedures, versioning, resource sharing, retention practices, janitor program utilization, and the like.

Janitor Program: An application which runs at scheduled intervals to manage business information by deleting, transferring, or archiving on-line data (such as email) which is at or past its scheduled active life. Janitor programs are sometimes referred to as “agents”—software that runs autonomously “behind the scenes” on user systems and servers to carry out business processes according to pre-defined rules. Janitor programs must include a facility to support disposition and process holds.

Java: Sun Microsystems’ Java is a platform-independent, programming language for adding animation and other actions to websites.

Jaz Drive: A removable disc drive. A Jaz drive holds up to 2 GB of data. Commonly used for backup storage as well as everyday use.

JMS: Jukebox Management Software.

Journal: A chronological record of data processing operations that may be used to reconstruct a previous or an updated version of a file. In database management systems, it is the record of all stored data items that have values changed as a result of processing and manipulation of the data.

Journaling: A function of e-mail systems (such as Microsoft Exchange and Lotus Notes) that copies sent and received items into a second information store for retention or preservation. Because Journaling takes place at the information store (server) level when the items are sent or received, rather than at the mailbox (client) level, some message-related metadata, such as user foldering (what folder the item is stored in within the recipient's mailbox) and the status of the "read" flag, is not retained in the journaled copy. The Journaling function stores items in the system's native format, unlike e-mail archiving solutions, which use proprietary storage formats that are designed to reduce the amount of storage space required. Journaling systems also lack the sophisticated search and retrieval capabilities contained in e-mail archiving solutions.

JPEG (Joint Photographic Experts Group): A compression algorithm for still images that is commonly used on the web.

Jukebox: A mass storage device that holds optical discs and loads them into a drive.

Juke-Box: Automated disc changer for high-performance, centralized storage for multifunction CD-ROM's & optical discs.

Jump Drive: See Key Drive.

Kerning: Adjusting the spacing between two letters.

Key Drive: A small removable data storage device that uses flash memory and connects via a USB port. Keydrives are also known as keychain drive, thumb drive, jump drive, USB flash drive. Can be imaged and may contain residual data.

Key Field: Database fields used for document searches and retrieval.

Keyword Search: A search of the text of documents in a database for documents containing one or more words that are specified by a user.

Keywords: Words designated by a user as important for searching purposes.

Kilobyte (KB): A unit of 1,024 bytes. See Byte.

Kofax Board: The generic term for a series of image processing boards manufactured by Kofax Imaging Processing. These are used between the scanner and the computer, and perform real-time image compression and decompression for faster image viewing, image enhancement, and corrections to the input to account for conditions such as document misalignment.

LAN (Local Area Network): A group of computers at a single location (usually an office or home) that are connected by phone lines or coaxial cable. See Network.

Landscape Mode: The image is represented on the page or monitor such that the width is greater than the height.

Laser Disc: Same as an optical CD, except 12" in diameter.

Laser Printing: A beam of light hits an electrically charged drum and causes a discharge at that point. Toner is then applied which sticks to the non-charged areas. Paper is pressed against the drum to form the image and is then heated to dry the toner. Used in laser printers and copying machines.

Latency: The time it takes to read a disc (or jukebox), including the time to physically position the media under the read/write head, seek the correct address and transfer it.

Latent Data: Latent or ambient data are deleted files and other data that are inaccessible without specialized forensic tools and techniques. Until overwritten, these data reside on media such as a hard drive in unused space and other areas available for data storage.

Leading/Ledding: The amount of space between lines of printed text.

Legacy Data, Legacy System: Legacy Data is information the development of which an organization may have invested significant resources and has retained its importance, but has been created or stored by the use of software and/or hardware that has become obsolete or replaced ("legacy systems"). Legacy data may be costly to restore or reconstruct when required for investigation or litigation analysis or discovery.

Legal Hold: A legal hold is a communication issued as a result of current or anticipated litigation, audit, government investigation or other such matter that suspends the normal disposition or processing of records. Legal holds can encompass business procedures affecting active data, including, but not limited to, backup tape recycling. The specific communication to business or IT organizations may also be called a "hold," "preservation order," "suspension order," "freeze notice," "hold order," or "hold notice."

Level Coding: Used in Bibliographical coding to facilitate different treatment, such as prioritization or more thorough extraction of data, for different categories of documents, such as by type or source.

LFP: IPRO Tech's image cross reference file; an ASCII delimited text file required for cross-reference of images to data.

Lifecycle: The records lifecycle is the life span of a record from its creation or receipt to its final disposition. It is usually described in three stages: creation, maintenance and use, and archive to final disposition.

Line Screen: The number of half-tone dots that can be printed per inch. As a general rule, newspapers print at 65 to 85 lpi.

Link: See Hyperlink.

Load file: A file that relates to a set of scanned images and indicates where individual pages belong together as documents. A load file may also contain data relevant to the individual documents, such as metadata, coded data and the like. Load files must be obtained and provided in prearranged formats to ensure transfer of accurate and usable images and data.

Local Area Network (LAN): See Network.

Logical File Space: The actual amount of space occupied by a file on a hard drive. The amount of logical file space differs from the physical file space because when a file is created on a computer, a sufficient number of clusters (physical file space) are assigned to contain the file. If the file (logical file space) is

not large enough to completely fill the assigned clusters (physical file space) then some unused space will exist within the physical file space.

Logical Volume: An area on the hard drive that has been formatted for files storage. A hard drive may contain a single or multiple volumes.

Lossless Compression: Exact construction of image, bit-by-bit, with no loss of information.

Lossy Compression: Reduces storage size of image by reducing the resolution and color fidelity while maintaining minimum acceptable standard for general use.

LTO (Linear Tape-Open): A type of backup tape which can hold as much as 400 GB of data, or 600 CDs depending on the data file format.

LZW (Lempel-Ziv & Welch): A common, lossless compression standard for computer graphics, used for most TIFF files. Typical compression ratios are 4/1.

Magenta: Used in four color printing. Reflects blue & red and absorbs green.

Magnetic/Optical Storage Media: Includes, but is not limited to, hard drives, backup tapes, CD-ROMs, DVD-ROMs, Jaz and Zip drives.

Magneto-Optical Drive: A drive that combines laser and magnetic technology to create high-capacity erasable storage.

Mailbox: An area on a storage device where email is placed. In email systems, each user has a private mailbox. When the user receives email, the mail system automatically puts it in the appropriate mailbox.

Make-Available Production: A process whereby what is usually a large universe of all potentially responsive documents are made available to the requestor; from this universe, the requestor then reviews and selects or tags the documents which they wish to obtain, and the producing party produces to the requestor only the selected documents. This is sometimes done under an agreement protecting against privilege and confidentiality waiver during the initial make available production; and the producing party, after the requestor has selected the documents they wish to obtain, reviews only the selected documents for privilege and confidentiality before the selected documents are physically produced to the requestor.

MAPI (Mail Application Program Interface): A Windows software standard that has become a popular email interface used by MS Exchange, GroupWise, and other email packages.

MAPI Mail Near-Line: Documents stored on optical discs or compact discs that are housed in the jukebox or CD changer and can be retrieved without human intervention.

Marginalia: Handwritten notes in the margin of the page in documents.

Master Boot Record: See Boot Sector.

Mastering: Making many copies of a disc from a single master disc.

MCA (Micro Channel Architecture): An IBM bus standard.

MDE (Magnetic Disc Emulation): Software that makes a jukebox look and operate like a hard-drive such that it will respond to all the I/O commands ordinarily sent to a hard drive.

MD5: message-digest algorithm meant for digital signature applications where a large message has to be “compressed” in a secure manner before being signed with the private key

Media: An object or device, such as a disc, tape, or other device, on which data is stored

Megabyte (M or MB): A unit of approximately 1 million bytes or 1024 KB. *See* Byte.

Memory: Data storage in the form of chips, or the actual chips used to hold data; “storage” is used to describe memory that exists on tapes or discs. *See* RAM and ROM.

Menu: A list of options, each of which performs a desired action such as choosing a command or applying a particular format to a part of a document.

Message Header: Message headers generally contain the identities of the author and recipients, the subject of the message, and the date the message was sent.

Metadata: Metadata is 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. *See also* Customer-Added Metadata, Document Metadata, Email Metadata, File System Metadata, User-Added Metadata and Vendor-Added Metadata. For a thorough discussion of Metadata, *see The Sedona Guidelines: Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age, Appendix D: Technical Appendix.*

Metadata Comparison: A method of de-duplication that compares file metadata and ignores content. *See* De-Duplication.

MFT (Master File Table): Index to files on a computer. If corrupt, a drive may be unusable, yet data may be retrievable using forensic methods.

MICR (Magnetic Ink Character Recognition): The process used by banks to encode checks.

Microfiche: Sheet microfilm (4” by 6”) containing reduced images of 270 pages or more in a grid pattern.

Migrated Data: Migrated Data is information that has been moved from one database or format to another.

Migration: Moving files to another computer application or platform; may require conversion to a different format.

Mirroring: The duplication of data for purposes of backup or to distribute Internet or network traffic among several servers with identical data. *See also* Disc Mirroring.

MIS: Management Information Systems.

MODEM: Modulator/Demodulator. A device translates digital data from a computer into analog signals (modulates) and transmits the information over telephones lines. Another modem at the receiving computer will receive the information, translate it back from analog to digital (demodulate) and store it.

Monochrome: Displays capable of only two colors, usually black and white, or black and green.

Mosaic: A web browser popular before the introduction of Netscape and Internet Explorer.

Mount, Mounting: The process of making off-line data available for on-line processing. For example, placing a magnetic tape in a drive and setting up the software to recognize or read that tape. The terms “load” and “loading” are often used in conjunction with, or synonymously with, “mount” and “mounting” (as in “mount and load a tape”). “Load” may also refer to the process of transferring data from mounted media to another media or to an on-line system.

MPEG-1 & -2: Two different standards for full motion video to digital compression/decompression techniques advanced by the Moving Pictures Experts Group. MPEG-1 compresses 30 frames/second of full-motion video down to about 1.5 Mbits/sec from several hundred megabytes. MPEG-2 compresses the same files down to about 3.0 Mbits/sec and provides better image quality.

MS-DOS: Microsoft (MS)-Disc Operating System. Used in PCs as the control system.

MTBF (Mean Time Between Failure): Average time between failures. Used to compute the reliability of devices/equipment.

MTTR (Mean Time To Repair): Average time to repair. The higher the number, the more costly and difficult to fix.

Multimedia: The combined use of different media; integrated video, audio, text and data graphics in digital form.

Multisynch: Analog video monitors which can receive a wide range of display resolutions, usually including TV (NTSC). Color analog monitors accept separate red, green & blue (RGB) signals.

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.

Natural Language Search: A manner of searching that permits the use of plain language without special connectors or precise terminology, such as “Where can I find information on William Shakespeare?” as opposed to formulating a search statement (such as “information” and “William Shakespeare”).

Near-Line Data: A term used to refer to data or a robotic storage device (robotic library) that houses removable media, uses robotic arms to access the media, and uses multiple read/write devices to store and retrieve records. Examples include optical discs.

Near-Line Data Storage: Storage in a system that is not a direct part of the network in daily use, but that can be accessed through the network. There is usually a small time lag between the request for data stored in near-line media and its being made available to an application or end-user. Making near-line data available will not require human intervention (as opposed to “off-line” data which can only be made available through human actions).

Network Gear: Refers to the actual hardware used in the operation of networks – for example routers, switches and hubs.

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.

Neural Network: Neural networks are made up of interconnected processing elements called units, which respond in parallel to a set of input signals given to each

NIST - National Institute of Standards and Technology - a federal technology agency that works with industry to develop and apply technology measurements and standards.

Node: Any device connected to a network. PCs, servers, and printers are all nodes on the network.

Non-Interlace: When each line of a video image is scanned separately. Computer monitors use non-interlaced video.

NOS (Network Operating System): See Operating System.

NSF: Lotus Notes Format Database File (i.e. database.nsf) Can be either an email database or the traditional type of fielded database.

Objects: In programming terminology, an object is a freestanding block of code that defines the properties of some thing. Objects are created and used in a high-level method of programming called object-oriented programming (OOP). OOP involves giving programming objects characteristics that can be transferred to, added to, and combined with other objects to make a complete program.

OCR (Optical Character Recognition): A technology process that translates and converts printed matter on an image into a format that a computer can manipulate (ASCII codes, for example) and, therefore, renders that matter text searchable. OCR software evaluates scanned data for shapes it recognizes as letters or numerals. All OCR systems include an optical scanner for reading text, and software for analyzing images. Most OCR systems use a combination of hardware (specialized circuit boards) and software to recognize characters, although some inexpensive systems operate entirely through software. Advanced OCR systems can read text in a large variety of fonts, but still have difficulty with handwritten text. OCR technology relies upon the quality of the imaged material, the conversion accuracy of the software, and the quality control process of the provider. The process is generally acknowledged to be only 80-85 percent accurate.

Official Record Owner: See Record Owner.

Off-Line Data: The storage of electronic data outside the network in daily use (e.g., on backup tapes) that is only accessible through the off-line storage system, not the network.

Off-Line Storage: Electronic records stored or archived on removable disc (optical, compact, etc.) or magnetic tape used for making disaster-recovery copies of records for which retrieval is unlikely. Accessibility to off-line media usually requires manual intervention and is much slower than on-line or near-line storage depending on the storage facility. The major difference between near-line data and offline data is that offline data lacks an intelligent disc subsystem, and is not connected to a computer, network, or any other readily-accessible system.

OLE: Object Linking and Embedding. A feature in Microsoft's Windows which allows each section of a compound document to call up its own editing tools or special display features. This allows for combining diverse elements in compound documents.

On-Line Review: The culling process produces a dataset of potentially responsive documents which are then reviewed for a final selection of relevant or responsive documents and assertion of privilege exception as appropriate. On-line Review enables the culled dataset to be accessed via PC or other terminal device via a local network or remotely via the Internet. Often, the On-Line Review process is facilitated by specialized software which provides additional features and functions which may include: collaborative access of multiple reviewers, security, user logging, search and retrieval, document coding, redaction, and privilege logging.

On-Line Storage: The storage of electronic data as fully accessible information in daily use on the network or elsewhere.

Online/On-Line: Connected (to a network).

Operating System (OS): An Operating system provides the software platform which directs the overall activity of a computer, network or system, and on which all other software programs and applications can run. In many ways, choice of an operating system will effect which applications can be run. Operating systems perform basic tasks, such as recognizing input from the keyboard, sending output to the display screen, keeping track of files and directories on the disc and controlling peripheral devices such as disc drives and printers. For large systems, the operating system has even greater responsibilities and powers - becoming a traffic cop to makes sure different programs and users running at the same time do not interfere with each other. The operating system is also responsible for security, ensuring that unauthorized users do not access the system. Examples of operating systems are UNIX, DOS, Windows, LINUX, Macintosh, and IBM's VM. Operating systems can be classified in a number of ways, including: multi-user (allows two or more users to run programs at the same time - some operating systems permit hundreds or even thousands of concurrent users); multiprocessing (supports running a program on more than one CPU); multitasking (allows more than one program to run concurrently); multithreading (allows different parts of a single program to run concurrently); and real time (instantly responds to input - general-purpose operating systems, such as DOS and UNIX, are not real-time).

Optical Discs: Computer media similar to a compact disc that cannot be rewritten. An optical drive uses a laser to read the stored data.

Optical Jukebox: See "Jukebox."

OST: A Microsoft Outlook information store that is used to save folder information that can be accessed offline.

Overwrite: To record or copy new data over existing data, as in when a file or directory is updated. Data that is overwritten cannot be retrieved.

PAB: A Microsoft Outlook list of recipients created and maintained by an individual user for personal use. The personal address book is a subset of the global address list (GAL).

PackBits: A compression scheme which originated with the Macintosh. Suitable only for black & white.

Packet: A unit of data sent across a network which may contain identify and routing information. When a large block of data is to be sent over a network, it is broken up into several packets, sent, and then reassembled at the other end. The exact layout of an individual packet is determined by the protocol being used.

Page File/Paging File: A file used to temporarily store code and data for programs that are currently running. This information is left in the swap file after the programs are terminated, and may be retrieved using forensic techniques. Also referred to as a swap file.

Page: A single image of the equivalent of “one piece of paper”. One or several pages make up a “Document.”

Parallel: Transmission of all the bits (e.g. in a character) at the same time. If the character has eight bits, there are eight wires. Faster and more expensive than serial where the eight bits would be sent, “sideways”, one at a time.

Partition: A partition is an individual section of computer storage media such as a hard drive. For example a single hard drive may be divided into several partitions. When a hard drive is divided into partitions, each partition is designated by a separate drive letter, i.e., C, D, etc.

Partition Table: The partition table indicates each logical volume contained on a disc and its location.

Partition Waste Space: After the boot sector of each volume or partition is written to a track, it is customary for the system to skip the rest of that track and begin the actual useable area of the volume on the next track. This results in unused or “wasted” space on that track where information can be hidden. This “wasted space” can only be viewed with a low level disc viewer. However, forensic techniques can be used to search these “wasted space” areas for hidden information.

Password: A secret code utilized, usually along with a user ID, in order to log on or gain access to a PC, network or other secure system, site or application.

Path: The hierarchical description of where a directory, folder, or file is located on a computer or network. In DOS and Windows systems, a path is a list of directories where the operating system looks for executable files if it is unable to find the file in the working directory. The list of directories can be specified with the PATH command. Path is also used to refer to a transmission channel, the path between two nodes of a network that a data communication follows, and the physical cabling that connects the nodes on a network.

Pattern Matching: A generic term that describes any process that compares one file’s content with another file’s content.

Pattern Recognition: Technology that searches data for like patterns and flags, and extracts the pertinent data, usually utilizing an algorithm. For instance, in looking for addresses, alpha characters followed by a comma and a space, followed by two capital alpha characters, followed by a space, followed by five or more digits, are usually the city, state and zip code. By programming the application to look for a pattern, the information can be electronically identified, extracted, or otherwise utilized or manipulated.

PCI: Peripheral Component Interface (Interconnect). A high-speed interconnect local bus used to support multimedia devices.

PCMCIA: Personal Computer Memory Card International Association. Plug-in cards for computers (usually portables), which extend the storage and/or functionality.

PDA (Personal Digital Assistant): A small, usually hand-held, computer which “assists” business tasks.

PDF (Portable Document Format): An imaging file format technology developed by Adobe Systems. PDF captures formatting information from a variety of applications in such a way that they can be viewed and printed as they were intended in their original application by practically any computer, on multiple platforms, regardless of the specific application in which the original was created. PDF files may be text-searchable or image-only. Adobe® Reader, a free application distributed by Adobe Systems, is required to view a file in PDF format. Adobe® Acrobat, an application marketed by Adobe Systems, is required to edit, capture text, or otherwise manipulate a file in PDF format.

Personal Computer (PC): computer based on a microprocessor and designed to be used by one person at a time

Petabyte (PB): A unit consisting of 1,000 or 1,024 terabytes. *See* Byte.

Phase Change: A method of storing information on rewritable optical discs.

Physical Ddisc: An actual piece of computer media, such as the hard disc or drive, floppy discs, CD-ROM discs, Zip discs, etc.

Physical File Space: When a file is created on a computer, a sufficient number of clusters (physical file space) are assigned to contain the file. If the file (logical file space) is not large enough to completely fill the assigned clusters (physical file space) then some unused space will exist within the physical file space. This unused space is referred to as file slack and can contain unused space, previously deleted/overwritten files or fragments thereof.

PICA: One sixth (1/6) of an inch. Used to measure graphics/fonts. There are 12 points per pica; 6 picas per inch; 72 points per inch.

Picture Element: The smallest addressable unit on a display screen. The higher the resolution (the more rows of columns), the more information can be displayed.

Pitch: Characters (or dots) per inch, measured horizontally.

PKI Digital Signature: A document or file may be digitally signed using a party’s private signature key, creating a “digital signature” that is stored with the document. Anyone can validate the signature on the document using the public key from the digital certificate issued to the signer. Validating the digital signature confirms who signed it, and ensures that no alterations have been made to the document since it was signed. Similarly, an email message may be digitally signed using commonly available client software that implements an open standard for this purpose, such as Secure Multipurpose Internet Mail Extensions (S/MIME). Validating the signature on the email can help the recipient know with confidence who sent it, and that it was not altered during transmission. *See* Certificate.

Plaintext: The least formatted and therefore most portable form of text for computerized documents.

Platter: One of several components that make up a computer hard drive. Platters are thin, rapidly rotating discs that have a set of read/write heads on both sides of each platter. Each platter is divided into a series of concentric rings called tracks. Each track is further divided into sections called sectors, and each sector is sub-divided into bytes.

PMS (Pantone Matching System): A color standard in printing.

POD (Print On Demand): Document images are stored in electronic format and are available to be quickly printed and in the exact quantity required, long or short runs.

Pointer: A pointer is an index entry in the directory of a disc (or other storage medium) that identifies the space on the disc in which an electronic document or piece of electronic data resides, thereby preventing that space from being overwritten by other data. In most cases, when an electronic document is “deleted,” the pointer is deleted, which allows the document to be overwritten, but the document is not actually erased.

Portable Volumes: A feature that facilitates the moving of large volumes of documents without requiring copying multiple files. Portable volumes enable individual CDs to be easily regrouped, detached and reattached to different databases for a broader information exchange.

Portrait Mode: A display where the height exceeds the width.

Preservation: The process of ensuring retention and protection from destruction or deletion all potentially relevant evidence, including electronic metadata. *See also* Spoliation.

Preservation Notice, Preservation Order: *See* Legal Hold.

Printout: A printed version of text of data, another term for which is hard copy.

Private Network: A network that is connected to the Internet but is isolated from the Internet with security measures allowing use of the network only by persons within the private network.

Privilege Data Set: The universe of documents identified as responsive and/or relevant, but withheld from production on the grounds of attorney-client privilege or work product.

Processing Data: In the context of this document, synonymous with Image Processing.

Production: The process of delivering to another party, or making available for that party’s review, documents deemed responsive to a discovery request.

Production Data Set: The universe of documents identified as responsive to document requests and not withheld on the grounds of attorney-client privilege or work product.

Production De-Duplication: Removal of a document if multiple copies of that document reside within the same production set. For example, if two identical documents are both marked responsive, non-privileged, production de-duplication ensures that only one of those documents is produced. *See* De-Duplication.

Production Number: Often referred to as the BATES number. A sequential number assigned to every page of a production for tracking and reference purposes. Often used in conjunction with a suffix or

prefix to identify the producing party, the litigation, or other relevant information. *See also* Bates Number.

Program: *See* Application and Software.

Properties: Fields of electronic information, or “metadata,” associated with a record or document such as creation date, author, date modified, blind cc’s and date received.

Protocol: Defines a common series of rules, signals and conventions that allow different kinds of computers and applications to communicate over a network. One of the most common protocols for networks is called TCP/IP.

Proximity Search: For text searches, the ability to look for words or phrases within a prescribed distance of another word or phrase, such as “accident” within 5 words of “tire.”

Public Network: A network that is part of the public Internet.

PST: A Microsoft Outlook email store. Multiple .pst files may exist and contain archived email.

QBIC (Query By Image Content): An IBM search system for stored images which allows the user to sketch an image, and then search the image files to find those which most closely match. The user can specify color and texture – such as “sandy beaches” or “clouds”.

Quality Control (QC): Steps taken to ensure that results of a given task, product or service are of sufficiently high quality; the operational techniques and activities that are used to fulfill requirements for quality. In document handling and management processes, this includes image quality (resolution, skew, speckle, legibility, etc.), and data quality (correct information in appropriate fields, validated data for dates, addresses, names/issues lists, etc.).

Quarter Inch Cartridge (QIC): Digital recording tape, 2000 feet long, with an uncompressed capacity of 5 GB.

Query: A request for specific information from a database or other data set.

Queue: A sequence of items such as packets or print jobs waiting to be processed. For example, a print queue holds files that are waiting to be printed.

RAID (Redundant Array of Independent Discs): A method of storing data on servers that usually combines multiple hard drives into one logical unit thereby increasing capacity, reliability and backup capability. RAID systems may vary in levels of redundancy, with no redundancy being a single, non-mirrored disc as level 0, two discs that mirror each other as level 1, on up, with level 5 being one of the most common. RAID systems are more complicated to copy and restore.

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. (*See* Dynamic Random Access Memory - DRAM).

Raster/Rasterized (Raster or Bitmap Drawing): A method of representing an image with a grid (or “map”) of dots. Typical raster file formats are GIF, JPEG, TIFF, PCX, BMP, etc.

Record: Information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business (ISO 15489(1)).

Record Custodian: A records custodian is an individual responsible for the physical storage and protection of records throughout their retention period. In the context of electronic records custodianship may not be a direct part of the records management function in all organizations. For example, some organizations may place this responsibility within their Information Technology Department, or they may assign responsibility for retaining and preserving records with individual employees.

Record Lifecycle: The time period from which a record is created until it is disposed.

Record Owner: The records owner is the subject matter expert on the contents of the record and is responsible for the lifecycle management of the record. This may be, but is not necessarily, the author of the record.

Record Series: A description of a particular set of records within a file plan. Each category has retention and disposition data associated with it, applied to all record folders and records within the category. (DOD 5015)

Record Submitter: The Record Submitter is the person who enters a record in an application or system. This may be, but is not necessarily, the author or the record owner.

Record: Information, regardless of medium or format that has value to an organization. Collectively the term is used to describe both documents and recorded data.

Records Hold: *See* Legal Hold.

Records Management: Records Management is the planning, controlling, directing, organizing, training, promoting, and other managerial activities involving the life-cycle of information, including creation, maintenance (use, storage, retrieval), and disposition, regardless of media.

Records Manager: The records manager is responsible for the implementation of a records management program in keeping with the policies and procedures that govern that program, including the identification, classification, handling and disposition of the organization’s records throughout their retention life. The physical storage and protection of records may be a component of this individual’s functions, but it may also be delegated to someone else. *See* Records Custodian.

Records Retention Period, Retention Period: The length of time a given records series must be kept, expressed as either a time period (*e.g.*, four years), an event or action (*e.g.*, audit), or a combination (*e.g.*, six months after audit).

Records Retention Schedule: A plan for the management of records listing types of records and how long they should be kept; the purpose is to provide continuing authority to dispose of or transfer records to historical archives.

Records Store: *See* Repository for Electronic Records.

Recover, Recovery: See Restore.

Redaction: A portion of an image or document is intentionally concealed to prevent disclosure of specific portions. Often done to avoid production of privileged or irrelevant materials.

Refresh Rate: The number of times per second a display (such as on a CRT or TV) is updated.

Region (of an image): An area of an image file that is selected for specialized processing. Also called a “zone.”

Registration: Lining up a forms image to determine which fields are where. Also, entering pages into a scanner such that they are correctly read.

Relative Path: An implied path.

Remote Access: The ability to access and use digital information from a location off-site from where the information is physically located. For example, to use a computer, modem, and some remote access software to connect to a network from a distant location.

Render Images: To take a native format electronic file and convert it to an image that appears as the original format file as if printed to paper.

Report: Formatted output of a system providing specific information.

Repository for Electronic Records: Repository for Electronic Records is a direct access device on which the electronic records and associated metadata are stored. (DoD 5015) Sometimes called a “records store” or “records archive.”

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.

Resolution: See DPI.

Restore: To transfer data from a backup medium (such as tapes) to an on-line system, often for the purpose of recovery from a problem, failure, or disaster. Restoration of archival media is the transfer of data from an archival store to an on-line system for the purposes of processing (such as query, analysis, extraction, or disposition of that data). Archival restoration of systems may require not only data restoration but also replication of the original hardware and software operating environment. Restoration of systems is often called “recovery”.

Retention Schedule: See Records Retention Schedule.

Reverse Engineering: The process of analyzing a system to identify its intricacies and their interrelationships, and create depictions of the system in another form or at a higher level. Reverse engineering is usually undertaken in order to redesign the system for better maintainability or to produce a copy of a system without utilizing the design from which it was originally produced. For example, one might take the executable code of a computer program, run it to study how it behaved with different input, and then attempt to write a program which behaved the same or better.

Review: The culling process produces a dataset of potentially responsive documents which are then examined and evaluated for a final selection of relevant or responsive documents and assertion of privilege exception as appropriate. *Also see* On-Line Review.

Rewriteable Technology: Storage devices where the data may be written more than once – typically hard drives, floppies and optical discs.

RFC822: Standard that specifies a syntax for text messages that are sent among computer users, within the framework of email.

RGB (Red, Green and Blue): The three primary colors in the additive color family which create all the computer color video signals for a computer's color terminal.

RIP: The procedures used to unbundle email collections into individual emails during the e-discovery process while preserving authenticity and ownership.

RIM: Records and information management.

RLE (Run Length Encoded): Compressed image format; supports only 256 colors; most effective on images with large areas of black or white.

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.

Root Directory: The top level in a hierarchical file system. For example on a PC, the root directory of your hard drive, usually C:, contains all the second-level subdirectories on that drive.

Rotary Camera: In microfilming, the papers are read "on the fly" with a camera that's synchronized to the motion.

Router: A device that forwards data packets along networks. A router is connected to at least two networks, commonly two LANs or WANs or a LAN and its ISP's network. Routers are located at gateways, the places where two or more networks connect.

Sampling Rate: The frequency at which analog signals are converted to digital values during digitization. The higher the rate, the more accurate the process.

Sampling: Sampling usually (but not always) refers to the process of testing a database for the existence or frequency of relevant information. It can be a useful technique in addressing a number of issues relating to litigation, including decisions about what repositories of data are appropriate to search in a particular litigation, and determinations of the validity and effectiveness of searches or other data extraction procedures.

SAN (Storage Area Network): A high-speed subnetwork of shared storage devices. A storage device is a machine that contains nothing but a disc or discs for storing data. A SAN's architecture works in a way that makes all storage devices available to all servers on a LAN or WAN. As more storage devices are added to a SAN, they too will be accessible from any server in the larger network. In this case, the server merely acts as a pathway between the end user and the stored data. Because stored data does not reside

directly on any of a network's servers, server power is utilized for business applications, and network capacity is released to the end user. *Also see* Network.

Scalability: The capacity of a system to expand without requiring major reconfiguration or re-entry of data. For example, multiple servers or additional storage can be easily added.

Scale-to-Gray: An option to display a black and white image file in an enhanced mode, making it easier to view. A scale-to-gray display uses gray shading to fill in gaps or jumps (known as aliasing) that occur when displaying an image file on a computer screen. Also known as grayscale.

Scanner: An input device commonly used to convert paper documents into images. Scanner devices are also available to scan microfilm and microfiche.

Scanning Software: Software that enables a scanner to deliver industry standard formats for images in a collection. Enables the use of OCR and coding of the images.

Schema: A set of rules or conceptual model for data structure and content, such as a description of the data content and relationships in a database.

Scroll Bar: The bar on the side or bottom of a window that allows the user to scroll up and down through the window's contents. Scroll bars have scroll arrows at both ends, and a scroll box, all of which can be used to scroll around the window.

SCSI (Small Computer System Interface): Pronounced "skuzzy." A common, industry standard, electronic interface (highway) between computers and peripherals, such as hard discs, CD-ROM drives and scanners. SCSI allows for up to 7 devices to be attached in a chain via cables. As of this writing, the current SCSI standard is "SCSI II," also known as "Fast SCSI."

SDLT (Super DLT): A type of backup tape which can hold up to 220 GB or 330 CDs, depending on the data file format. *See* DLT.

Search: *See* Compliance Search, Concept Search, Contextual Search, Boolean Search, Full-Text Search, Fuzzy Search, Index, Keyword Search, Pattern Recognition, Proximity Search, QBIC, Sampling, *and* Search Engine.

Search Engine: A program that enables search for keywords or phrases, such as on web pages throughout the World Wide Web.

Sector: A sector is normally the smallest individually addressable unit of information stored on a hard drive platter, and usually holds 512 bytes of information. Sectors are numbered sequentially starting with 1 on each individual track. Thus, Track 0, Sector 1 and Track 5, Sector 1 refer to different sectors on the same hard drive. The first PC Hard discs typically held 17 sectors per track. Today, they can hold thousands of sectors per track.

Serial Line Internet Protocol (SLIP): A connection to the Internet in which the interface software runs in the local computer, rather than the Internet's.

Serif: The little cross bars or curls at the end of strokes on certain type fonts.

Server: Any central computer on a network that contains data or applications shared by multiple users of the network on their client PCs. 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. *See* File Server.

Service-Level Agreement: A service-level agreement is a contract that defines the technical support or business parameters that a service provider or outsourcing firm will provide its clients. The agreement typically spells out measures for performance and consequences for failure.

SGML/HyTime: A multimedia extension to SGML, sponsored by DOD.

SHA-1: Secure Hash Algorithm, for computing a condensed representation of a message or a data file specified by FIPS PUB 180-1.

Signature: *See* Certificate.

SIMM (Single, In-Line Memory Module): A mechanical package (with “legs”) used to attach memory chips to printed circuit boards.

Simplex: One-sided page(s)

Skewed: Tilted images. *See* De-skewing.

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.

Smart Card: A credit card size device which contains a microprocessor, memory and a battery.

SMTP (Simple Mail Transfer Protocol): The protocol widely implemented on the Internet for exchanging email messages.

Software application: *See* Application and Software.

Software: Any set of coded instructions (programs) stored on computer-readable media that tells a computer what to do. Includes operating systems and software applications.

Speckle: Imperfections in an image as a result of scanning paper documents that do not appear on the original. *See* De-speckling.

Splatter: Data that should be kept on one disc of a jukebox goes instead to multiple platters.

Spoliation: Spoliation is the destruction of records which may be relevant to ongoing or anticipated litigation, government investigation or audit. Courts differ in their interpretation of the level of intent required before sanctions may be warranted. *See The Sedona Guidelines: Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age, Guideline 3.*

SPP (Standard Parallel Port): *See* Centronics.

SQL (Structured Query Language): a standard fourth generation programming language (4GL - a programming language that is closer to natural language and easier to work with than a high-level language). The popular standard for running database searches (queries) and reports.

Stand-Alone Computer: A personal computer that is not connected to any other computer or network, except possibly through a modem.

Standard Generalized Markup Language (SGML): An informal industry standard for open systems document management which specifies the data encoding of a document's format and content.

Status Bar: A bar at the bottom of a window that is used to indicate the status of a task. For example, when an email message is sent, the status bar will fill with dots indicating that a message is being sent.

Steganography: The hiding of information within a more obvious kind of communication. Although not widely used, digital steganography involves the hiding of data inside a sound or image file. Steganalysis is the process of detecting steganography by looking at variances between bit patterns and unusually large file sizes.

Storage Device: A device capable of storing data. The term usually refers to mass storage devices, such as disc and tape drives.

Storage Media: See Magnetic or Optical Storage Media.

Subjective Coding: The coding of a document using legal interpretation as the data that fills a field, versus objective data that is readily apparent from the face of the document, such as date, type, author, addresses, recipients and names mentioned. Usually performed by paralegals or other trained legal personnel.

Subtractive Colors: Since the colors of objects are white light *minus* the color absorbed by the object, they are called subtractive. This is how ink on paper works. The subtractive colors of process ink are CMYK (Cyan, Magenta, Yellow and Black) and are specifically balanced to match additive colors (RGB).

Suspension Notice, Suspension Order: See Legal Hold.

SVGA (Super Video Graphics Adapter): A graphics adapter one which exceeds the minimum VGA standard of 640 by 480 by 16 colors. Can reach 1600 by 1280 by 256 colors.

Swap File: A file used to temporarily store code and data for programs that are currently running. This information is left in the swap file after the programs are terminated, and may be retrieved using forensic techniques. Also referred to as a page file or paging file.

System Administrator (sysadmin, sysop): The person in charge of keeping a network working.

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.

T1: A high speed, high bandwidth leased line connection to the Internet. T1 connections deliver information at 1.544 megabits per second.

T3: A high speed, high bandwidth leased line connection to the Internet. T3 connections deliver information at 44.746 megabits per second.

Tape Drive: A hardware device used to store or backup electronic data on a magnetic tape. Tape drives are usually used to back up large quantities of data due to their large capacity and cheap cost relative to other data storage options.

Taxonomy: The science of categorization, or classification, of things based on a predetermined system. In reference to Web sites and portals, a site's taxonomy is the way it organizes its data into categories and subcategories, sometimes displayed in a site map.

TCP/IP (Transmission Control Protocol/Internet Protocol): A collection of protocols that define the basic workings of the features of the Internet.

Telephony: Converting sounds into electronic signals for transmission.

Templates, Document: Sets of index fields for documents, providing framework for preparation.

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 process or spreadsheet applications.

Terabyte: A unit of 1,000 or 1,024 gigabytes, or approximately a trillion bytes.

TGA: Targa format. This is a "scanned format" – widely used for color-scanned materials (24-bit) as well as by various "paint" and desktop publishing packages.

Thin Client: A networked user computer that acts only as a terminal and stores no applications or user files. May have little or no hard drive space. *See Client*.

Thread: A series of postings on a particular topic. Threads can be a series of bulletin board messages (for example, when someone posts a question and others reply with answers or additional queries on the same topic). A thread can also apply to chats, where multiple conversation threads may exist simultaneously.

Thumb Drive: *See Key Drive*.

Thumbnail: A miniature representation of a page or item for quick overviews to provide a general idea of the structure, content and appearance of a document. A thumbnail program may be standalone or part of a desktop publishing or graphics program. Thumbnails take considerable time to generate, but provide a convenient way to browse through multiple images before retrieving the one needed. Programs often allow clicking on the thumbnail to retrieve it.

TIFF (Tagged Image File Format): One of the most widely used and supported graphic file formats for storing bit-mapped images, with many different compression formats and resolutions. File name has .TIF extension. Can be black and white, gray-scaled, or color. Images are stored in tagged fields, and programs use the tags to accept or ignore fields, depending on the application. The format originated in the early 1980's.

TIFF Group III (compression): A one-dimensional compression format for storing black and white images that is utilized by many fax machines. *See* TIFF.

TIFF Group IV (compression): A two-dimensional compression format for storing black and white images. Typically compresses at a 20-to-1 ratio for standard business documents. *See* TIFF.

Toggle: A switch that is either on or off, and reverses to the opposite when selected.

Tool Kit Without An Interesting Name (TWAIN): A universal toolkit with standard hardware/software drivers for multi-media peripheral devices.

Toolbar: The row of buttons right below the menu that perform special functions quickly and easily.

Topology: The geometric arrangement of a computer system. Common topologies include a bus (network topology in which nodes are connected to a single cable with terminators at each end), star (local area network designed in the shape of a star, where all end points are connected to one central switching device, or hub), and ring (network topology in which nodes are connected in a closed loop; no terminators are required because there are no unconnected ends). Star networks are easier to manage than ring topology.

Track: Each of the series of concentric rings contained on a hard drive platter.

True Resolution: The “true” optical resolution of a scanner is the number of pixels per inch (without any software enhancements).

Typeface: There are over 10,000 typefaces available for computers. The general categories are: oldstyle (faces have slanted serifs, gradual thick to thin strokes and a slanted stress - the “O” appears slanted), modern (faces have thin, horizontal serifs, radical thick to thin strokes and a vertical stress - the “O” does not appear to slant); slab serif (faces have thick, horizontal serifs, little or no thick-to-thin in the strokes and a vertical stress - the “O” appears vertical); sans serif (faces have no serifs), script (from elaborate handwriting styles to casual, freeform, unconnected letter forms), decorative unusual fonts (designed to be very different and attention getting).

Ultrafiche: Microfiche which can hold 1,000 documents/sheet as opposed to the normal 270.

UMS: Universal messaging system.

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.

Unitization – Physical and Logical: The assembly of individually scanned pages into documents. Physical Unitization utilizes actual objects such as staples, paper clips and folders to determine pages that belong together as documents for archival and retrieval purposes. Logical unitization is the process of human review of each individual page in an image collection using logical cues to determine pages that belong together as documents. Such cues can be consecutive page numbering, report titles, similar headers and footers and other logical indicators. This process should also capture document relationships, such as parent and child attachments. *See also* Attachment.

UNIX: A software operating system.

Upgrade: new or better version of some hardware or software.

Upload: To send a file from one computer to another via modem, network, or serial cable. With a modem-based communications link, the process generally involves the requesting computer instructing the remote computer to prepare to receive the file on its disc and wait for the transmission to begin.

URI (Uniform Resource Indicators): a URL is a URI.

URL (Uniform Resource Locators): the addressing system used in the World Wide Web and other Internet resources. The URL contains information about the method of access, the server to be accessed and the path of any file to be accessed. A URL looks like this: http://thesedonaconference.org/publications_html.

URL: See Address.

User-Added Metadata: Data or work product created by a user while reviewing a document, including annotations and subjective coding information.

V.32bis: The ITU (*see ITU*) standard for 14.4 kbs modem communications.

V.34: The proposed ITU (*see ITU*) standard for 28.8 kbs modem communications.

Validate: In the context of this document, confirm or ensure well grounded logic, and true and accurate determinations.

VAR/VAD/VASD: Value-Added Reseller/Value-Added Dealer/Value-Added Specialty Distributor. Companies or people who sell computer hardware or software *and* “add-value” in the process. Usually, the value added is specific technical or marketing knowledge and/or experience.

VDT (Video Display Terminal): generic name for all display terminals.

Vector: Representation of graphic images by mathematical formulas. For instance, a circle is defined by a specific position and radius.

Vendor-Added Metadata: Data created and maintained by the electronic discovery vendor as a result of processing the document. While some vendor-added metadata has direct value to customers, much of it is used for process reporting, chain of custody and data accountability. Contrast with Customer-Added Metadata.

Verbatim Coding: Extracting data from documents in a collection in a way that matches exactly as the information appears in the documents.

Version, Record Version: A particular form or variation of an earlier or original record. For electronic records the variations may include changes to file format, metadata or content.

Vertical De- Duplication: A process through which duplicate data are eliminated within a single custodial or production data set. *See* Content Comparison, File level Binary Comparison Horizontal De- duplication, Meta Data Comparison.

VGA (Video Graphics Adapter): A PC industry standard, first introduced by IBM in 1987, for color video displays. The *minimum* dot (pixel) display is 640 by 480 by 16 colors. Then “Super VGA” was introduced at 800 x 600 x 16, then 256 colors. VGA can extend to 1024 by 768 by 256 colors. Replaces EGA, an earlier standard and the even older CGA. Newer standard displays can range up to 1600 by 1280.

Video Electronics Standards Association (VESA): concentrates on computer video standards.

Video Scanner Interface: A type of device used to connect scanners with computers. Scanners with this interface require a scanner control board designed by Kofax, Xionics or Dunord.

Virus: a self-replicating program that spreads by inserting copies of itself into other executable code or documents. A program into which a virus has inserted itself is said to be infected, and the infected file (or executable code that is not part of a file) is a host. Viruses are a kind of malware (malicious software). Viruses can be intentionally destructive, for example by destroying data, but many viruses are merely annoying. Some viruses have a delayed payload, sometimes referred to a bomb. The primary downside of viruses is uncontrolled self-reproduction, which desecrates or engulfs computer resources.

Vital Record: A record that is essential to the organization’s operation or to the reestablishment of the organization after a disaster.

VoIP (Voice over Internet Protocol): Telephonic capability across an IP connection; increasingly used in place of standard telephone systems.

Volume: A volume is a specific amount of storage space on computer storage media such as hard drives, floppy discs, CD-ROM discs, etc. In some instances, computer media may contain more than one volume, while in others, one volume may be contained on more than one disc.

Volume Boot Sector - When a partition is formatted to create a volume, a volume boot sector is created to store information about the volume. One volume contains the operating system and its volume boot sector contains code used to load the operating system when the computer is booted up.

VPN (Virtual Private Network): A secure network that is constructed by using public wires to connect nodes. For example, there are a number of systems that enable creation of networks using the Internet as the medium for transporting data. These systems use encryption and other security mechanisms to ensure that only authorized users can access the network and that the data cannot be intercepted.

WAV: *File extension name for Windows sound files.* “.WAV” files can reach 5 Megabytes for one minute of audio.

Web Site: A collection of Uniform Resource Indicators (URIs), including Uniform Resource Locators (URLs), in the control of one administrative entity. May include different types of URIs (*e.g.*, file transfer protocol sites, telnet sites, as well as World Wide Web sites). *See* URI and URL.

Workflow, Ad Hoc: A simple manual process by which documents can be moved around a multi-user review system on an “as-needed” basis.

Workflow, Rule-Based: A programmed series of automated steps that route documents to various users on a multi-user review system.

Workgroup: A group of computer users connected to share individual talents and resources as well as computer hardware and software – often to accomplish a team goal.

WORM Discs: Write Once Read Many Discs. A popular archival storage media during the 1980s. Acknowledged as the first optical discs, they are primarily used to store archives of data that cannot be altered. WORM discs are created by standalone PCs and cannot be used on the network, unlike CD-ROM discs.

WORM (Write-Once, Read-Many): Data storage devices (e.g. CD-ROM's) where the space on the discs can *only* be written *once*. The data is *permanently* stored. This is often today's primary media for archival information. Common disc sizes run from 5.25" (1.3 gigabytes) to 12" (8 to 10 gigabytes) capacities. There is also a 14" disc (13 to 15 gigabytes), only manufactured by Kodak's optical storage group. WORMs can also be configured into jukeboxes. There are various technologies. The expected viable lifetime of a WORM is at least 50 years. Since it's impossible to change, the government treats it just like paper or microfilm and it is accepted in litigation and other record-keeping applications. On the negative side, there is no current standard for how WORMs are written. The only ISO standard is for the 14" version, manufactured only by one vendor. A 5.25" standard is emerging from the European Computer Manufacturing Association but is not yet accepted. Further, WORM discs are written on both sides, but there are currently no drives that read both sides at the same time. As for speed, WORM is faster than tape or CD-ROM, but slower than magnetic. Typical disc access times run between 40 and 150 milliseconds (compared with 11 ms for fast magnetic discs and 300 ms for CD-ROM). Data transfer rates run between 1 and 2 MB/sec (compared with 5 to 10 for magnetic discs and 600KB/sec for CD-ROM).

WWW (World Wide Web): All of the computers on the Internet which use HTML-capable software (Netscape, Explorer, etc.) to exchange data. Data exchange on the WWW is characterized by easy-to-use graphical interfaces, hypertext links, images, and sound. Today the WWW has become synonymous with the Internet, although technically it is really just one component.

WYSIWYG: "What You See Is What You Get" – Display and software technology which shows on the computer screen exactly what will print. Often requires a large, high-density monitor.

X.25: A standard protocol for data communications.

XML: See Extensible Markup Language.

Zip Drives: A floppy disc drive that can hold a large amount of data, usually as much as 750 megabytes or more. Often used for backing up hard discs.

ZIP: A common file compression format that allows quick and easy storage for transport.

Zone OCR: An add-on feature of the imaging software that populates document templates by reading certain regions or zones of a document, and then placing the text into a document index.

- END -