

“Help Me Help You” with E-Discovery: Working With Outside Counsel to Prepare for Discovery of Electronically Stored Information

Gordon, Arata, McCollam, Duplantis & Eagan, L.L.P.
2009 Duck Lunch
Nina Wessel English and Steven W. Copley
(With invaluable assistance from Matthew R. Emmons)

I. Common E-Discovery Terms

What is reprinted below is a non-copyrighted list of common e-discovery terms from **ELaw Exchange**, “Form or Forms,” available at <http://www.elawexchange.com>.

The following are the common "form or forms" of "electronically stored information."

- Native File
- Metadata
- Database
- Spreadsheet
- Image
- Text, ASCII, and Conversion Formats
- Video and Audio
- Paper
- Automated Litigation Support (ALS) Form and Online ESI Depository
- Forensic Image

1. Native file

Computer application software programs generally maintain their data in proprietary file formats commonly referred to as “native file” format. A native file is the file format used by a specific software application. These files often have a specific file extension, which are the last three or more letters following the file name and period, such as “smith.PST” for a Microsoft Outlook file (PST) relating to a person named “Smith.” These proprietary data file formats determine which software program can read the data file and it also contains the file’s metadata. Unless otherwise stipulated, an agreement to exchange "native" files will include the entire computer file, including all metadata.

Reference Sources:

- *Arkfeld on Electronic Discovery and Evidence*, § 3.4, Directories, Files and File Formats; 3.6(B), Classification of Data; 5.3, ESI Forms and Disclosure Formats; 7.7(G), ESI Form(s); 7.7(G)(2), Native File — Default Format (for legal considerations regarding native files.).
 - *Williams v. Sprint/United Mgmt. Co.*, 230 F.R.D. 640, 656-657 (D. Kan. 2005). The Court ordered an employer in an employment discrimination case to restore the metadata it had “scrubbed” or “erased” from Excel spreadsheet files and “unlock” them. The Court stated, “[w]hen the Court orders a party to produce an electronic document in the form in which it is regularly maintained, i.e., in its native format or as an active file, that production must include all metadata”
-

2. Metadata

The Court in *Aguilar v. Immigration & Customs Enforcement Div.*, No. 07-8224, 2008 U.S. Dist. LEXIS 97018, at *11-15 (S.D.N.Y. Nov. 20, 2008) stated:

1. Types of Metadata

Metadata, frequently referred to as “data about data,” is electronically-stored evidence that describes the “history, tracking, or management of an electronic document.” It includes the “hidden text, formatting codes, formulae, and other information associated” with an electronic document. . . . *Autotech Techs. Ltd. P’Ship v. AutomationDirect.com, Inc.*, 248 F.R.D. 556, 557 n.1 (N.D. Ill. 2008) (Metadata includes “all of the contextual, processing, and use information needed to identify and certify the scope, authenticity, and integrity of active or archival electronic information or records”). Although metadata often is lumped into one generic category, there are at least several distinct types, including substantive (or application) metadata, system metadata, and embedded metadata. . . .

There are three types of “metadata.”

a. Substantive Metadata

Substantive metadata, also known as application metadata, is “created as a function of the application software used to create the document or file” and reflects substantive changes made by the user. (citation omitted). This category of metadata reflects modifications to a document, such as prior edits or editorial comments, and includes data that instructs the computer how to display the fonts and spacing in a document. (citation omitted)

Substantive metadata is embedded in the document it describes and remains with the document when it is moved or copied. . . .

b. System Metadata

System metadata “reflects information created by the user or by the organization's information management system.” (citation omitted) This data may not be embedded within the file it describes, but can usually be easily retrieved from whatever operating system is in use. . . . Examples of system metadata include data concerning “the author, date and time of creation, and the date a document was modified.” . . . Courts have commented that most system (and substantive) metadata lacks evidentiary value because it is not relevant. (citations omitted) System metadata is relevant, however, if the authenticity of a document is questioned or if establishing “who received what information and when” is important to the claims or defenses of a party. (citation omitted) This type of metadata also makes electronic documents more functional because it significantly improves a party's ability to access, search, and sort large numbers of documents efficiently. . . .

c. Embedded Metadata

Embedded metadata consists of “text, numbers, content, data, or other information that is directly or indirectly inputted into a [n]ative [f]ile by a user and which is not typically visible to the user viewing the output display” of the native file. . . . Examples include spreadsheet formulas, hidden columns, externally or internally linked files (such as sound files), hyperlinks, references and fields, and database information. . . . This type of metadata is often crucial to understanding an electronic document. For instance, a complicated spreadsheet may be difficult to comprehend without the ability to view the formulas underlying the output in each cell. . . .

3. Database

In computer terminology, a database is simply a collection of mutually related data or information stored in computer record fields. They are organized collections of information similar to index cards, phone books, or file cabinets of documents. In business, all kinds of data, from e-mail and contact information to financial data and records of sales, are stored in some form of a database. Databases track employee information, payrolls, job classifications, retirement benefits and a host of other business related information. The Fed. R. Civ. P. 34, Advisory Committee Note of 2006 recognized “dynamic databases” and how databases may store different forms of ESI, “[e]lectronically stored information may exist in dynamic databases and other forms far different from fixed expression on paper. . . . Using current technology, for example, a party might be called upon to produce word processing documents, email messages, electronic spreadsheets, different image or sound files, and material from databases.” A subset of a database can be produced by the use of queries and reports. They can be exported into a comma, delimited, spreadsheet or other file format.

Reference Sources:

- *Arkfeld on Electronic Discovery and Evidence*, §§ 3.11 Structure and Type of Electronic Information, 5.3, ESI Forms and Disclosure Formats, and 7.7(G), ESI Form(s).
- *N.A.A.C.P. v. Acusport Corp.*, 210 F.R.D., 268, 278-279 (E.D.N.Y. Sept. 18, 2002) (case provides extensive discussion of the different components of a database)
- *Cook v. Deloitte & Touche, LLP*, No. 03-3926, 2005 U.S. Dist. LEXIS 22252, at *8, 14 (D.N.Y. Sept. 30, 2005) (database contained a file for each employee reflecting each telephone call or letter between [employer] and the employee, including notes summarizing the substance of the contact . . . [and a database] describing vacant positions).
- *In re Seroquel Prods. Liab. Litig.*, MDL 1769, 2007 U.S. Dist. LEXIS 5877, at *8 (D. Fla. Jan. 26, 2007) (Databases holding information about customers, production, employee performance, internal processes, etc.)

4. Spreadsheet

Spreadsheet application programs perform simple and complex mathematical calculations automatically. Spreadsheets can be used in a variety of business functions, and oftentimes are used by individuals to keep their financial and other records. Spreadsheets have long been used in business and are of interest for electronic discovery because of their content. Calculations and mathematical analyses, mailing lists, to-do lists, attendance rosters, invoices, real estate closing statement calculations, truth in lending statements and mortgage payments, loan calculations and amortization schedules are a few of the uses of spreadsheets. Database information can be downloaded into a spreadsheet format.

"Native" spreadsheet computer files contain:

Calculations or formulas that are not visible in a printout version (only the result of the calculation is visible);

Hidden cells, columns, rows and post-it style comments;

Hidden worksheets;

Hidden formulas; and

Display of all rows and columns.

Reference Sources:

- *Arkfeld on Electronic Discovery and Evidence*, § 3.12 Structure and Type of Electronic Information and § 7.7(G), ESI Form(s)
- *Public Citizen v. Carlin*, 2 F. Supp.2d 1, 14 (D.D.C. 1997). (Paper print-outs of computer spreadsheets only display the results of calculations made on the spreadsheet, while the actual electronic version of the spreadsheet will show the formula used to make the calculations.)
- *Williams v. Sprint/United Mgmt. Co.*, 230 F.R.D. 640 (D. Kan. 2005) (The court ordered an employer in an employment discrimination case to restore the metadata it had scrubbed or erased from Excel spreadsheet files and unlock them.)

5. Image

Imaging is a process where paper or ESI is scanned into a system and stored electronically as a “picture.” These digitized computer files of documents are known as “images.” Images may be searchable or non-searchable. Metadata from the original document is generally not converted by the producing party if you convert and print a file to a TIFF or PDF format. ESI can be converted to images for disclosure purposes, redaction or presentation in the courtroom. In addition, images may be linked to a database to assist in the retrieval process.

If an image is not searchable it can be converted into searchable text by using optical character recognition (OCR). OCR is a process of converting letters or numbers that appear on an image or printed page to a bit mapped image and then into ASCII that can be searched. However, the process is not 100% accurate and it may be expensive to “clean up” the converted image data. An image that is created from ESI is generally searchable since the ESI is converted directly into an image with the accompanying “text.” This is usually accomplished by using the PDF image format. In addition, ESI can be converted to a TIFF nonsearchable image and be accompanied with searchable “text.”

The two most popular image forms for litigation are TIFF and PDF.

TIFF (Tagged Image File Format). A TIFF is a standard proprietary file format for storing images as bit maps. It is used especially for scanning documents because it can support any size, resolution and color depth. TIFF images are not searchable. If the document is OCR’ed, which is not 100% accurate, then it will become partially searchable.

PDF (Portable Document Format). PDF stands for Portable Document Format and is also a proprietary file format (Adobe, Inc.) that preserves the fonts, images, graphics and layouts of any source document, such as a Microsoft Word document, regardless of the application and platform used to create it. After converting a file to a PDF format anyone with (a free) Adobe Reader software can view the document as it originally appeared in the application program. This precludes the necessity of having to obtain a licensed copy of the application program to view the document. There are two different types of PDF files:

IMAGE ONLY format is an exact electronic picture of the paper document. It cannot be word searched, unless the image is subsequently OCR'ed; IMAGE FORMAT with SEARCHABLE TEXT FORMAT is an electronic picture or image of the document that also contains background "hidden" text that can be word searched using Adobe Reader software. Documents, spreadsheets, e-mail and graphics can all be converted to TIFF or PDF.

Reference Sources:

- *Arkfeld on Electronic Discovery and Evidence*, § 5.3 (D), ESI Forms and Disclosure Formats and § 7.7(G), ESI Form(s)

Fed. R. Civ. P. 34(a) states: "(a) Scope. Any party may serve on any other party a request (1) to produce . . . documents or electronically stored information — including . . . images . . . stored in any medium from which information can be obtained (emphasis added)."

- The Fed. R. Civ. P. 34, Advisory Committee Note of 2006 introduced new terminology for the ESI "image" form and stated, "[i]mages, for example, might be hard-copy documents or electronically stored information."
- *Williams v. Sprint/United Mgmt. Co.*, 230 F.R.D. 640, 643, n.8 (D. Kan. 2005) ("TIFF (Tagged Image File Format) is one of the most widely used and supported graphic file formats for storing bit-mapped images, with many different compression formats and resolutions. A TIFF file is characterized by its '.tif' file name extension. (citation omitted)").
- *In re Payment Card Interchange Fee & Merch. Disc. Antitrust Litig.*, No. 05-1720, 2007 U.S. Dist. LEXIS 2650, at *6,7,15 (E.D.N.Y. Jan. 12, 2007). (TIFF images without searchable text are an unacceptable form of production and a violation of the Rules Advisory Committee Proviso that data ordinarily kept in electronically searchable form should not be produced in a form that removes or significantly degrades this feature.)

- *CP Solutions PTE, Ltd. v. GE*, No. 04-2150, 2006 U.S. Dist. Lexis 27053, at *1-15 (D. Conn. Feb. 6, 2006)

6. Text, ASCII, and Conversion Formats

Text or ASCII. “Text” or “ASCII” (American Standard Code for Information Interchange) documents are those documents that have the “text” of a document stored in a computer file. These documents can be word or phrase searched and one can instantly access the exact location of the words in the text documents. ASCII is a format that most computer programs recognize for transferring data between programs and to conduct “text” searches. Essentially any document produced is in a “text” format, if it is in an ASCII format. Once in an ASCII format, it can be imported and searched in a search and retrieval software program. An example of a “text” document is the deposition of a witness. Other examples of “text” documents include; business documents, trial transcripts, witness interviews, expert reports, etc.

Conversion Formats. Generally, in order to transfer data between different database, spreadsheet and automated litigation support (ALS) programs you have to convert data into a format that is recognized by both programs. For example, one common format for transferring data from one application to another is “comma-delimited” in which each piece of data is separated by a “comma.” Most database and spreadsheet programs are able to import and export “comma-delimited data.” Any character can be used to separate the data, but the common separators or delimiters are the comma (usually referred to as CSV (comma-separated value)), text, vertical bar, space and the tab key. Column headers are usually included as the first line and used as “field descriptors” for identification purposes.

Reference Sources:

- *Zakre v. Norddeutsche Landesbank Girozentrale*, No. 03-257, 2004 U.S. Dist. LEXIS 6026, at *1-2 (D.N.Y. April 9, 2004) (plaintiff is able to search the ESI provided in a text-searchable format “for single words or phrases, or combinations of words or phrases”).
- *In re Payment Card Interchange Fee & Merch. Disc. Antitrust Litig.*, No. 05-1720, 2007 U.S. Dist. LEXIS 2650, at *12-14 (D.N.Y. Jan. 12, 2007) (“‘OCR’ refers to ‘optical character recognition,’ a computer software program that translates images of text into a format that can be searched or ‘read’ electronically.”).
- *Pace v. Int’l Mill Serv.*, No. 05-69, 2007 U.S. Dist. LEXIS 34104, at *1-2 (D. Ind. May 7, 2007). In a pre-amendment case, the Court denied plaintiff’s motion to compel and noted that plaintiff’s

consultant alleged that the defendant provided “a DVD that contained ‘Microsoft Office Excel comma separated value files,’ rather than PDF files. . . . [and] that the data . . . was missing field descriptions and was provided in comma-separated, rather than the requested text-delineated format. The lack of field descriptions was corrected by [the defendant].”

- *J.C. Assocs. v. Fid. & Guar. Ins. Co., No. 01-2437*, 2006 U.S. Dist. LEXIS 32919 (D.D.C. May 25, 2006). The Court ordered the plaintiff to make available to the defendant an OCR-scan program (OmniPage Pro) to permit the defendant to convert and search these files for specific keywords, check for privilege and then disclose to the plaintiff.

7. Paper

Fed. R. Civ. P. 34, Advisory Committee Note of 2006 recognizes that: “[t]he form of production is more important to the exchange of electronically stored information than of hard-copy materials, although a party might specify hard copy as the requested form.”

Reference Sources:

- *In re Bristol-Myers Squibb Sec. Litig.*, 205 F.R.D. 437, 443-444 (D.N.J. 2002) The Court noted that, “[o]f course, in some instances, paper, rather than electronic, production may still be the preferable method of discovery.”
- *MacNamara v. City of New York*, No. 04-9216, 2006 U.S. Dist. LEXIS 82926, at *16-17 (D.N.Y. Nov. 13, 2006). The defendants produced electronic arrest records of the plaintiffs and others to which the plaintiffs objected. The plaintiffs contended that “significant errors, edits and omissions” occurred at the data entry stage for arrest records and requested “the handwritten worksheets.” The Court agreed and ordered the city “to produce . . . Worksheets for non-party arrestees as well as named plaintiffs, subject to the ‘attorneys’-eyes-only’ designation”

8. Automated Litigation Support (ALS) Form and Online ESI Depository

Technically a disclosure of ESI in an automated litigation support (ALS) format is not a “form.” However, many litigants, courts and agencies will agree or order that the exchange of ESI be provided in an ALS format. For example, the Court may require that all ESI be provided in a Summation “load” file which will enable the requesting party to immediately load, search, analyze and produce ESI reports in Summation. These “load” files may contain a database, images and text. Essentially, ESI is preprocessed to be used immediately with ALS systems.

Automated litigation support (ALS) generally refers to computer operations that support legal functions in litigation. Summation and Concordance are two of the leading litigation ALS systems. An ALS system can manage a large volume of documents, electronic data, transcripts and other data in a secure environment for quick retrieval and analysis during litigation. In addition, online ESI depositories, similar to such as Lextranet or CaseVault, are web based ALS systems and are increasingly being used for hosting and management of ESI.

Reference Sources:

- *Arkfeld on Electronic Discovery and Evidence*, § 5.2(B) Automated Litigation Support System and § 5.2(C), Online ESI Depository
- *O’Bar v. Lowe’s Home Ctrs., Inc.*, No. 04-00019, 2007 U.S. Dist. LEXIS 32497, *21 (D.N.C. May 2, 2007) (“If load files were created in the process of converting Native Files to Static Images, or if load files may be created without undue burden or cost, load files should be produced together with Static Images.”).
- *Quinby v. WestLB AG*, No. 04-7406, 2006 U.S. Dist. LEXIS 64531, at *17-18 (D.N.Y. Sept. 5, 2006). The Court denied defendant’s request to shift the costs of reformatting data from a “.tif” to a “.dii” format because the defendant failed to raise the issue in its initial motion. “[n9 ‘.tif’ or ‘Tagged Image Format’ is a commonly used electronic format for digitized pictures of documents. It is often used to produce e-mails for loading onto Summation, a software program designed to assist attorneys in searching, organizing and analyzing documents . . .] [n10 ‘.dii’ is a format which facilitates the loading of files onto Summation and may be created from .tif files.]”

9. Audio and Video

Audio has moved from analog recording with LPs (long-playing records) and tape cassettes as the playback medium to digital recording using computers with a digital sound playback. Digital audio files are usually compressed for less storage requirements and faster transmission. The most popular audio file format today is MP3. MP3 is a standard technology and format for compressing a sound sequence into a very small file while preserving the original level of sound quality when it is played.

Video generally refers to recording, manipulating and displaying moving images in a format that can be presented on a television or on a computer monitor. It is a recording produced with a video recorder (camcorder) or some other device that captures full motion.

Reference Sources:

- *Arkfeld on Electronic Discovery and Evidence*, § 3.18, Video; § 3.20, Audio

II. Preparation for the Rule 26(f) Conference

- Federal Rule of Civil Procedure 26(f) requires that the parties meet and confer at least 21 days before a scheduling conference and consider a discovery plan. The discovery plan must state the parties' views and proposals “on any issues about disclosure or discovery of electronically stored information, including the form or forms in which it should be produced.” Fed. R. Civ. Pro. 26(f)(3)(C).

The following items should be addressed with your outside counsel prior to receiving discovery requests and prior to the Rule 26(f) conference.

1. **Review your document retention plan:** Do you have a document retention plan that addresses documents and ESI? Is it being applied? Are all employees on notice of the terms of the plan? Are employees following the plan?
2. **Prepare and send litigation hold notices:** When litigation is “reasonably anticipated,” must cease normal document retention policy. Notice must interrupt the usual document retention/destruction policies to preserve any documents and ESI that could be at issue in the litigation (or anticipated litigation). In addition to suspending the usual document/ESI destruction procedures, the hold must be sent to all relevant employees and must instruct them to turn off any auto-delete functions.
3. **Meet with key employees to take stock of all documents and ESI:** Include your outside counsel in meetings with key employees. You should question all employees as to what devices they use to transact business (i.e. in addition to work computers, some might be using a BlackBerry, private email accounts, home computers, etc.).
4. **Meet with your IT personnel:** Include your outside counsel in this meeting so that outside counsel can hear first-hand how the company stores data, deletes data, recovers data and so that outside counsel can learn about the network architecture for the company. Outside counsel will need to be able to explain your company’s systems and all of the searches performed by the company should there be a Motion to Compel.
5. **Determine ease/difficulty and cost of producing ESI.** Depending on the complexity of your ESI, this might be a good time to hire an outside consultant.
6. **Preserve backup tapes:** If relevant, remove backup tapes from the rotation to prevent them from being recycled. Note: Most courts have held that backup tapes are not “readily accessible” particularly when they are used for purposes of disaster recovery. Backup tapes are “accessible” if they are used for information retrieval. Discovery can be had for non-readily accessible ESI only with a showing of need.
7. **Determine if any relevant documents/information have been destroyed.** You are protected from claims of spoliation by Fed. Rule Civ. P. 37(e)’s safe harbor provision

- so long as you took necessary steps to implement a litigation hold process to preserve relevant ESI upon learning of anticipated litigation.
8. **Accommodate outside counsel’s monitoring.** Outside counsel can no longer issue a one-time litigation hold to the client and pass on responsibility to the client to comply with that hold. Rather, courts are consistently requiring outside counsel to reissue, update and amend the litigation hold as needed.
 9. **Decide what you will request at Rule 26(f) conference.** The conference could be used to limit what the parties will request (i.e. parties will not request texts, voicemails, thumb drives, backup tapes, word searches, key employee searches by name, etc.). Decide in what format you want the ESI produced (native format, tiff, pdf, paper). (Note: Rule 34(b)(2)(E)(ii) says that if a request does not specify a form for producing ESI, it must be produced in the form (1) in which “it is ordinarily maintained,” (i.e. Native Format) or (2) in “a reasonably usable form.”)
 10. **Decide who will attend the Rule 26(f) conference.** Consider having in-house IT personnel attend and suggesting the same to the other side. Or consider hiring an outside IT consultant to attend.
 11. **Consider drafting a “clawback agreement”.** Agreement should include affirmative statement that there is no substantive waiver of the attorney client privilege. (Note: Even with a clawback agreement, parties must still conduct a privilege review to avoid waiving the privilege.)
 12. **Consider hiring outside IT consultant?**
 - a. Consultant could help you/counsel determine the most efficient way to produce ESI.
 - b. Consultant could help you/counsel determine in what format you would like to request the information.
 - c. Consultant could help you/counsel determine what information is readily accessible (and therefore must be produced) and what is not.

Sources: Kenneth J. Withers, *Electronically Stored Information: The December 2006 Amendments to the Federal Rules of Civil Procedure*, 4 Nw. J. Tech. & Intell. Prop. 171 (2006), <http://www.law.northwestern.edu/journals/njtip/v4/n2/3>; Merrill Corporation, www.merrillcorp.com/law; *Things to Think About for Your Rule 26(f) Meeting*, available at <http://bowtielaw.wordpress.com>.

III. Suggested Reading

A. Significant Cases Involving E-Discovery

- *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309 (S.D.N.Y. 2003) ("Zubulake I")
- *Zubulake v. UBS Warburg LLC*, 216 F.R.D. 280 (S.D.N.Y. 2003) ("Zubulake III")
- *Zubulake v. UBS Warburg LLC*, 220 F.R.D. 212 (S.D.N.Y. 2003) ("Zubulake IV")
- *Zubulake v. UBS Warburg, LLC*, 229 F.R.D. 422 (S.D.N.Y. 2004) ("Zubulake V")
- *Zubulake v. UBS Warburg LLC*, 382 F.Supp.2d 536 (S.D.N.Y. 2005) ("Zubulake VI")
- *Coleman (Parent) Holdings, Inc. v. Morgan Stanley & Co., Inc.*, 2005 WL 679071 (Fla. Cir. Ct. Mar. 1, 2005)
- *Qualcomm Inc. v. Broadcom Corp.*, 2008 WL 66932 (S.D. Cal. Jan. 7, 2008)

B. Blogs & Websites

- www.elawexchange.com
- www.pdfforallawyers.com
- www.bowtielaw.wordpress.com
- www.electronicdiscoveryblog.com
- www.arkfeld.globs.com
- www.ediscoverylaw.com
- www.docnativeblog.wordpress.com

C. Law Review Article

- *Electronically Stored Information: The December 2006 Amendments of the Federal Rules of Civil Procedure*, Kenneth J. Withers, available at <http://is.gd/jPKV>