



2016 CLM Annual Conference
April 6-8, 2016
Orlando, FL

Reduce Cost and Risk during Discovery

E-DISCOVERY GLOSSARY

Understanding e-discovery definitions and concepts is critical to working with vendors, opposing counsel, litigation support teams and others involved in requesting and producing documents and electronically stored information (ESI) for litigation. Across the EDRM there are terms unique to the discovery process. Here you will find some of the most important.

Accessibility

Under the “two-tier rule” established under FRCP 26(b)(2)(B), a party need not provide discovery of ESI from sources that the party identifies as not reasonably accessible because of undue burden or cost; on motion to compel discovery or for a protective order, the party from whom discovery is sought must show that the information is not reasonably accessible because of undue burden or cost.

Active Data

Data that is directly available to operating system and/or application software.

Admissibility

The admissibility of electronic records is still evolving as most organizations strive to move from paper to paperless records. However, the admissibility of electronic information is somewhat more complex, raising issues as to the methodology used in data collection and the chain of custody of the electronic data once it has been collected.

Amended Federal Rules of Civil Procedure

See “Federal Rules of Civil Procedure.”

Analysis

The process of determining relevancy of electronic discovery materials through evaluation based on the variables of the case.

Archival Data

Digital information that is retained for long-term storage, not immediately available and often stored on removable media.

Archive

A copy of data on a computer drive or on a portion of a drive, maintained for historical reference.

Attachment

A memorandum, letter, spreadsheet or any other electronic document appended to another document or email.

Backup

A copy of active data, intended for use in restoration of data.

Backup Tape

Magnetic tape used to store copies of ESI for restoration or recovery purposes.

Bates Number

Sequential numbering used to track documents and images in data sets. Each page has a unique number.

Boolean Search

Boolean refers to a system of logic developed by an early computer pioneer, George Boole. In Boolean searching, an “and” operator between two words results in a search for documents containing both of the words. An “or” operator between two words creates a search for documents containing either of the target words. A “not” operator between two words creates a search result containing the first word but excluding the second.

Chain of Custody

Documentation regarding possession, movement and location of evidence from the time it was obtained to the time it is presented in court. Failure to maintain a complete chain of custody may result in the inadmissibility of electronic information.

Clawback Agreement

An agreement that sets forth procedures to protect against waiver of privilege due to inadvertent production of documents or data.

Collection

The process of gathering electronically stored information.

Computer Forensics

Specialized techniques to recover, authenticate and analyze electronic data.

Concept Search

Analyzing conceptual groups of words in a document to understand the true meaning, rather than searching only for a word (keyword).

Container File

One file that contains multiple documents and document types. Requires decompression or ripping to process.

Contextual Search

Searching surrounding text to analyze the context in which a word is used.

Cost Shifting

Shifting the cost or a portion of the cost of production of inaccessible electronically stored documents to the requesting party. Understanding where your information lives and the information governance policies and procedures in place within your organization is critical to successful arguments for cost shifting or cost sharing. Discovery requests may cover email, databases, voice mail, instant messaging systems and other proprietary applications. These systems are rapidly evolving and do not represent the same burden for all companies when making arguments for cost shifting. Implementing legal holds or producing records from traditional voice mail systems take considerable time and money; newer unified messaging systems often make responding to the same request relatively easy. Similarly, restoring records from traditional tape backup systems can be time- and cost-intensive; near-line storage, by contrast, does not present the same challenges.

Cost Shifting

Shifting the cost or a portion of the cost of production of inaccessible electronically stored documents to the requesting party.

Culling

Removing a document prior to production or review; generally reduces the volume of data that is produced or reviewed.

Custodian

See “Data Custodian”

Custodian Deduplication

Culls a document if multiple copies of that document reside within the same custodian’s data set. For example, if Mr. A and Mr. B each have a copy of a specific document and Mr. C has two copies, the system will maintain one copy each for Mr. A, Mr. B and Mr. C. Contrast with case deduplication and production deduplication.

Data Formats

The organization of information for display, storage or printing. Data is maintained in certain common formats so that it can be used by various programs, which may only work with data in a particular format. This term is commonly used in the industry when asking another person about the state in which particular information exists. For example, “What format is it in, PDF or HTML?”

Data Mining

The process to cull data to extract ESI for production.

Data Production

Electronic evidence must be delivered to multiple parties involved in a legal matter, including opposing counsel, partner firms, requesting government agencies and others. Depending on the recipients, you may require different delivery formats.

De-NIST

Screening files against the NIST list of computer file types. Separates those files generated by a user from those generated by a system.

Deduplication

The process of identifying (or some vendors include actually removing) additional copies of identical documents in a document collection. There are three types of de-duplication: case, custodian and production.

Discovery Conference

See "Meet and Confer."

Discovery Cost Allocation

The distribution of the costs incurred by organizations who are compelled to produce ESI.

Discovery Response

Activities performed in response to a request for discovery.

Document Metadata

Data stored in the document about the document. Often this data is not immediately viewable in the software application used to create/edit the document, but can be accessed via a "Properties" view. Contrast with file system metadata and email metadata.

E-Discovery

See "Electronic Discovery."

Early Case Assessment (ECA)

A preliminary process of estimating the risk, cost and time required to handle discovery requests for electronically stored information (ESI) in a legal case. ECA allows attorneys to review client data without the cost or time of full processing. A good ECA tool will allow an attorney to run search terms, cull data and view many reports of the data set.

Electronic Discovery

The process of identifying, preserving, collecting, processing, reviewing and producing electronically stored information (ESI) for legal review.

Electronic Discovery Reference Model

The Electronic Discovery Reference Model is a framework to describe the phases of activities around electronic discovery (www.EDRM.net).

Email Metadata

Data stored in the email about the email. Often this data is not even viewable in the email client application used to create the email. The amount of email metadata available for a particular email varies greatly depending on the email system. Contrast with file system metadata and document metadata.

Embedded Metadata

Text, numbers, content, data or information that is directly or indirectly input into a native file by a user and which is not typically visible to the user viewing the output of the display of the native file on screen or as a printout.

Form 35 (FRCP)

Report of the parties' planning meeting in which the parties jointly propose the agreed upon discovery plan.

Hash/Hash Coding

Algorithm that represents a unique value, like a digital fingerprint. The process of creating a unique algorithm that is unique to every document.

Identification

A phase of the electronic discovery process involving the identification of all relevant sources of electronically stored information.

Imaged Copy

A "mirror image" bit-by-bit copy of a hard drive, i.e., a complete replication of the physical drive regardless of how the drive is organized or whether the image created contains meaningful data in whole or in part. From an imaged copy of a hard drive it is possible to reconstruct the entire contents and organization of the source drive from which it was taken.

Legal Hold/Preservation

A legal hold is an essential element of a company's overall records management program, particularly when it comes to electronic information. It needs to be issued to demonstrate a company's good faith and reasonable effort to comply with its discovery obligations. The reality, however, is that the full implications of the legal hold process may not be fully understood by all parties, particularly outside the legal department. Proactive coordination and planning among corporate counsel, outside counsel, IT and other key stakeholders are imperative to ensure good faith compliance in the face of anticipated litigation.

Meet and Confer (FRCP Rule 26(f))

A rule within the FRCP that requires parties to meet prior to a scheduling conference in federal court to discuss and agree upon discovery of information and evidence relevant to the case.

Meet and Confer 26(f) Conference

The meet and confer conference for electronic discovery has moved from a nice-to-have to a requirement under the amended Federal Rules of Civil Procedure. A major component of preparing for a 26(f) meet and confer conference is a "map" of the litigant's ESI content: where it is, what it is, how to preserve it, how to collect it, etc. This defensive requirement can be turned into a strategic advantage when counsel is well informed as to the location and nature of ESI, as well as the costs necessary to produce it.

Metadata

Handling and processing electronic evidence present new and unique challenges that are vastly different from working with traditional printed documents. Unique among these is the handling of metadata – document attributes as to creation, modification, authorship and potentially more details based on the application used to create the electronic documents. Metadata may be used for admissibility purposes, demonstrating the chain of custody for a particular piece of ESI. It may also be used in the preservation and review processes, identifying information to be held and facilitating the culling of duplicate documents, respectively. Simply opening a file or copying it to another location may modify the hidden metadata. In order to prevent spoliation, proper methods must be used in the collection and review of electronic documents. Similarly, metadata that represents privileged information must be carefully removed prior to production. This can be a complex process requiring deep technical expertise and experience.

Native File/Format

The source document, as collected from the source computer or server, before any conversion or processing of the document.

Native Format Review (also Native Review)

Reviewing ESI using the software used to create it originally. For example, using Microsoft Word in the review process to open/review a .DOC (MS Word Document format) file.

Near Deduplication

The elimination of electives with "near duplicate" similarities, i.e., a document that was sent to multiple custodians.

NIST

National Institute of Standards and Technology.

OCR

Optical Character Recognition, a method of translating printed text and images into a form that a computer can manipulate (into ASCII codes, for example). An OCR system enables you to scan a printed document directly into a computer file.

Presentation

Presentation of the preserved, collected, processed, reviewed, analyzed and produced ESI at a legal proceeding.

Preservation

The process of retaining and protecting all relevant evidence from destruction or deletion.

Privilege

Technology, when properly applied, has the ability to rapidly increase the rate of productivity exponentially. However, when improperly utilized, small mistakes can have large-scale effects – particularly when dealing with the inadvertent disclosure of privileged documents. Clawback and quick-peek agreements may help buffer the risk, but are only partially effective. And without proper steps to maintain privilege, courts may find it has been unintentionally waived.

Processing

Capturing an electronic data image or a representation of the image, generally in native format, entering it into a computer system and manipulating it so that it can be exported into a review application.

Production

To electronically deliver ESI to a variety of recipients or for use in other systems.

Quick Peek

ESI is made available to opposing party before being reviewed for privilege, confidentiality or privacy. Strict guidelines are required to prevent waiver.

Record Custodian

Person responsible for the storage and protection of records throughout the record retention period.

Review

Examination of potentially relevant data sets, or ESI, for relevancy, privilege and confidentiality in advance of production.

Rule 16 (FRCP)

Pretrial conference – FRCP Rule 16 may provide a party with an opportunity to discuss settlement without giving the appearance of having initiated the conversation.

Rule 26 (FRCP)

General provisions governing discovery; duty of disclosure.

Rule 37 (FRCP)

FRCP 37(e), formerly 37(f), provides a safe harbor when data is lost or overwritten in the normal course of business.

Rule 502 (FRE)

The proposed Federal Evidence Rule 502 is intended to reduce the risk of forfeiting the attorney-client privilege or work product protection “so that parties need not scrutinize production of documents to the same extent as they do now.”

Sampling

Testing a database or ESI to determine the frequency of relevant information.

Sanctions

The preservation of ESI can be costly for many large organizations. Over-preservation may result in escalating costs as information is produced at an exponential rate. Failure to preserve enough can result in a wide range of possible penalties, including monetary, issue, expert and case-related sanctions. Companies need to develop defensible processes that strike the balance between preservation and business needs. Such policies must also be effective and supportive of good faith efforts to identify and preserve potentially responsive ESI.

Spoilation

While FRCP 37(f) provides a safe harbor for routine, good faith deletion of e-discovery, FRCP 37 also provides for sanctions where the producing party fails to provide e-discovery outside of the safe harbor. In addition to sanctions, spoliation of ESI may result in an adverse inference, an award of attorneys' fees and possibly an adverse judgment. These risks can be mitigated through documented good faith efforts to preserve potentially responsive ESI.

Technology-Assisted Review (TAR)

(also called Computer Assisted Review and Predictive Coding) – The use of machine learning technologies to categorize an entire collection of documents as responsive or nonresponsive, based on human review of a subset of the document collection. These technologies typically rank the documents from “most likely” to “least likely” to be responsive to a specific information request. This ranking can then be used to “cut” or partition the documents into one or more categories, such as potentially responsive or not, in need of further review or not, and so on.

TIFF

Tagged Image File Format – a graphic file format used for storing still-image bitmaps. TIFFs are stored in tagged fields and programs use the tags to accept or ignore fields, depending on the application.