An Introduction to e-Discovery.

You've heard the term, but what does it mean?

Through our introductory series, we will de-mystify e-Discovery. We will go step by step over ten weeks and cover every aspect of the process. We will cover all the key issues, and all the key questions you need to ask. So when it comes time to make the important decisions, you'll have a solid understanding of what's at stake.

What is e-Discovery?

E-Discovery is short for Electronic Discovery, and deals with the management of Electronically Stored Information (ESI) in the discovery stage of litigation. A corollary to e-Discovery is Electronic Records Management (ERM), which is the practice used to manage ESI in a non-litigation setting. For the purpose of the this series, we will use E-Discovery and ERM interchangeably. In 2006 the Federal Rules of Civil Procedure (FRCP) were amended to include rules how litigants should go about handling ESI in civil litigation. The FRCP were further amended in 2015 to further clarify how the FRCP would deal with e-Discovery. The specific e-Discovery rules relate to FRCP 16, 26, 33, 34, 37 and 45. These rules generally cover:

1. A requirement for parties meet and confer in order to develop a discovery plan.
2. Obligations to preserve ESI;
3. Safe Harbor provisions for inadvertently lost ESI; and
4. Procedure for recalling ESI that was inadvertently produced

Why do you need to know about it?

Failure to follow the ESI rules outlined in FRCP could result in sanctions for law firms, spoliation claims against clients and waiver of privilege claims if not adequately addressed. Discovery in litigation is inherently legal in nature, and as such, courts have held that law firms share an obligation with their clients to preserve electronic records. Courts have also held that law firms were jointly and severally liable for monetary sanctions for failure to advise clients about litigation holds and their obligations to preserve relevant ESI. For clients, failure to adhere to e-Discovery rules could lead to a default judgment against them.

Trustpoint's Summer Series is designed as an introductory primer for ESI related topics. For a more in-depth discussion of e-Discovery and TrustPoint One's litigation solutions, please feel free to send any questions to Michael.Harris@Trustpoint.One.

eDiscovery Boot Camp.

Week 1 - Meet & Confer

At the start of any matter requiring discovery, the FRCP 26(f) mandates that both parties make an effort to meet and resolve any issues or disputes. This responsibility should be taken seriously as the court is required to impose monetary sanctions on any party who fails to meet and confer in good faith.

Required Parties

Attorneys and relevant parties that are involved in the matter or are aware of where the ESI data resides.

In advance of this meeting, a series of interviews should be conducted internally with the client or the client's IT department to determine the computer networks and client's electronic devices that host the ESI relevant to the matter, including the company's current data retention policies, any previous electronic storage method as well as current IT structure for email, cloud, and relevant mobile devices.

Considerations

Some of the largest and most common issues arise because the right people are not at the meeting. It is important to make sure that the everyone involved be present. This includes the review project manager, the ESI project manager (tech team handling the data) as well as attorneys from both sides that will be intimately engaged in the process.

Agenda

This required meeting provides an opportunity for all parties involved in the matter to discuss relevant discovery issues related to the case. Involved parties have the freedom to be frank and open about the data and data systems, providing a sense of timeline regarding when data can be collected, processed and reviewed before producing to the other side.

- While discussing the agenda - you'll want to create a tentative schedule for deadlines and future meetings to insure that the plan set out in this meeting is being executed as intended.
- Create a workflow to handle any changes or disagreements should they arise.
- Discuss possible cost saving measures, sometimes providing clarity for the more important data.

eDiscovery Boot Camp.

Week 2 - Legal Holds / Custodian Identification

The legal hold process is generally the most visible part of eDiscovery. It is easy to understand but success often lies in the details. For example, it is common for there to be multiple legal holds in place at one time, and making sure that you are aware of all relevant data locations and possible cross-border privacy rules can quickly complicate things.

FRCP 30(b)(6)

Allows a party to serve notice on an organization instead of an individual. This deposition notice requires the organization to designate individual(s) that are qualified to answer questions on the identified topics to be covered in the deposition.

Custodian Interviews

Create a questionnaire that gives you information regarding your clients involvement in the case matter. A basic set of questions regarding custodian's role in the company, location of desk, computers/software used, types of files received, or any other method of receiving and sending information regarding the case. This includes questions involving where data has been stored, intricate case details and changes/ history during the specific time period.

Legal Hold

Once a party can reasonably anticipate litigation, it must suspend its routine document retention/destruction policy and institute a "legal hold" to ensure document preservation.

Implementation

As part of the custodian identification process, the interested party will want to verify that there is no data spoliation in the discovery process. Topics of concern include:

- Legal hold policy that was enacted in response to the matter and all steps that were taken within the organization to implement this policy across the required custodians.
- Substance of the communication regarding legal hold policy implementation, timeline of issuance and parties involved.
- Current policies regarding auto-deletion/targeted back-ups of data and process involved in disabling them.
- Process for insuring compliance and tracking any problems that may arise in understanding its implementation.

Considerations

For large companies, managing legal holds can be its own full time job. While the process is most often carried out by outside counsel, it can be managed by eDiscovery consultants who specialize in the nuanced management of data retention.

eDiscovery Boot Camp.

Week 3 - Data Collection

Data Collection is often the most technical and complex phase of eDiscovery. It requires minimal disruption to the client's business but needs to preserve data integrity and chain of custody. It needs to be comprehensive without being over-inclusive. This phase is often considered just a task for the IT deptartment, but we have found a combination of legal and technical experience is essential.

Process

As part of an efficient collection, a detailed plan and chain of custody regarding data handling will need to be established. As discussed in last week's preservation, the actual method of how data should be preserved is often left to the parties that need to produce data. Unlike preservation, data collection is very process oriented and requires an exercise in technical dexterity and organization.

Metadata

Metadata is information about information. It is who created the document, who modified it, and when. In short, it provides additional background information to the data that is being collected and reviewed. It is one of the more complex topics related to eDiscovery, as it depends on technology's own ability to capture it.

Data Format

As part of collecting data, plans should include considerations for how the data should be collected.

Data Storage

Given the many changes in tech in the past few decades, one thing to note in setting up a data collection workflow is to get a firm understanding of data priority and storage. This generally involves answers from custodians hard drives, network drives and email. Generally a client's IT group will be able to start this process the fastest, as they are most familiar with where the data is stored.

Active – Data that is being used on a regular basis, often involves:
Cloud – Clients are moving toward cloud based tech for storage solutions, especially as more of the employees tend to work remotely. This also includes any data stored on mobile devices, cloud based applications or social media platforms.
Storage/Offline – For many clients, data related to the matter can reside in an offline server or backup drives/tapes that use uncommon compression tools which lead to a more difficult collection.

Considerations

As part of a successful collection plan, special attention must be given to the scope, size, timeline and order of the data being collected.

Add questions that we would ask?

Performing the Collection

Once the data has been identified, the next step is to consider who will physically collect the data, keeping in mind the level of expertise and cost.

Self Collection by the employee.
Outlook makes it easy for custodians to create PSTs of their emails and copy relevant shared and private drives. However, this is often the riskiest option as it can be done incorrectly and questioned by the court as to whether this is a defensible practice.

Client Collection by internal IT Dept.

External Collection by Vendor
Includes collection on premises or remote collection performed by the Vendor.

Additional Considerations

Chain of Custody
Hash code/digital fingerprinting data
Audit Trail of collections

eDiscovery Boot Camp.

Week 4 - Search Terms/Assessing Data

The right search terms make all the difference in sifting through data. You don't want to be caught unaware, be reactive, and spend more than you should because of issues with ESI. We have found that our clients are much more prepared and stay within budget by spending extra time analyzing their data and using effective search terms.

Narrowing the Scope

Employees often have multiple roles within a corporation which leads to a lot of useless data during collection. One of the tried and true methods in culling data for review is through relevant time periods and search terms. Creating a good set of search terms is essential to the discovery process and taking the time to assess and validate the list is one of the easiest ways to keep costs down. Knowledgeable consultants can also help by maximizing the efficiency of using the correct search syntax.

The Search Terms

Every day the sheer volume of data generated by an individual grows at an exponential rate. With rigid deadlines, one of the simplest forms of defensible practice is the utilization of keyword searches.

- During the meet & confer stage, create a timeline to discuss agreed upon search terms.
- Search technology. Having a list of search terms doesn't guarantee success in collection. Functionality between how the terms should be created for a vendor's processing platform vs client's internal business system should be considered.
- Collaboration: It is important that legal team work closely with the discovery consultant and tech team to verify that the operators and syntax used are correct.

Creating the Terms

There are a number of ways that can be used to create the terms and phrases. Taking the time to craft an accurate list that is comprehensive without being over inclusive is an important fact in reducing costs.

- **Pleadings/Subpoena** – The issue being investigated and relevant parties are a good starting source for keywords. In addition, factual allegations or focus of the investigation are also great resources for the initial set of search terms.
- **Client** – Discussing the crux of the litigation with the client should also provide some insight regarding the common terminology and issues that may be easily searchable.
- **ESI** – Analyzing the data itself may provide clues into additional terms that should be added to the list. Looking at critical custodians' file names, file folders, and project names in their emails can also be used.
- **Opposing or co-counsel** - Involving other parties to create a search term list may be helpful in adding the final set of terms. While the other methods contribute to fact-finding investigations, opposing counsel may add terms that will provide terms related to legal issues that include state of mind and knowledge of the allegation.

Timing

Search terms can be applied at various times in the collection and processing process. Choosing the right method is often dictated by the nature of matter.

Collection phase. Many end clients with IT teams prefer to run the search terms themselves. While this is generally the least expensive solution, it can also lead to some of the costliest delays down the road. If the review phase reveals any incorrect application of search terms, a second collection can cause costly delays in rushed processing and review.

Processing Phase. Search terms are run against the data during processing. This is usually the best compromise as the cost of ingesting the data for processing is low compared to the cost of injesting data post processing. Running the terms in this way allows you to get an accurate sense of volume and a good way to validate your collected data.

Post Processing in Database. Data is collected and processed without being run against search terms. This will provide a much larger volume of processed data. It is generally the most expensive method but it does provide the fastest way to adapt to constantly changing priorities, as you have all your data available to be run against any number of different searches.

Strategies

With every search that's performed, data should be assessed to see if the relevant hits and doc counts match up with expectations. If not- search terms may either be too broad or too narrow.

Attorneys should spend time with custodians to craft an original set of terms that would capture as much of the data as possible. This also includes words and phrases that may be in different languages.

After the creation of the first part of the list, it helps to work with a discovery consultant to determine if additional changes need to be made. During this process, it is important that the consultant is provided enough information regarding the history and intent of any terms/phrases that are not easily understandable. This allows any changes to provide the same results and conform to the tech requirements.

Search terms should be tested and re-evaluated if the results are either too large or too small. In addition, take time to sample the results to make sure that the retrieved data set actually includes what it should. Keep a record of all changes to the term list and document reasons for the final selection.

eDiscovery Boot Camp.

Week 5 - ECA Analysis/Data Processing

ECA is a preliminary analysis that provides a high level view of the data to determine if the collected information meets the requirements of the scope of discovery. This phase is also a good time to estimate eDiscovery costs for the client and validate the collected data.

Early Case Assesment

Early Case Assessment is a term that encompasses a defensible strategy to collect, analyze and streamline the processing and review of data.

Ingestion – (sometimes called processing). Raw data (native file metadata and textual content) is taken and converted into a standard, normalized format. This process also includes the handling of various types of files: such as password protected/encrypted documents, imaging (B&W/color), OCRing requirements as well as language detection.

There is a duty to the client to execute a defensible process in keeping a chain of custody of the data and avoiding spoliation and corruption during processing.

This phase is critical in pre-planning for the review and QC process. If the volume of data is considered too large for review after processing, you still have time to figure out what else to do.

Assessment.

Before processing any data, it is important to take a step back and make sure that the team has a proper strategy in place to check for any potential failures. To insure that this process goes smoothly and avoids the costs of re-collection, there are important things to consider:

- Confirmation and agreements establishing the processing protocols, acceptable risks, de-duplication, and culling.
- Quality Control methods to validate error-free processing and post-processing.
- Data types, formats and electronic or physical delivery.
- Complications – is some of the data very old? Possible corruption due to physical damage?
- Conversion of container files, cataloguing and itemization of files and folders.
- Standard methods of reporting and auditing including chain of custody during each phase of data handling.
- Security requirements regarding encryption, physical storage of the data, and personnel access.
- Scheduling for substantial compliance deadlines and custodian availability.
- Content analysis and de-duplication and indexing.

Processing

Once the data has been properly isolated and collected, it is now ready to be processed. This step often requires an iterative approach, depending on whether the results yield the information you expect.

Analysis of the data source. Often data is presented in container files for ease of delivery and to address cybersecurity concerns.

- Data size – How large is the data overall? Per custodian? This is usually one of the easiest and earliest indicators that provides a sense of volume, cost, and timeline.
- Metadata – These fields are generally preserved in collection and many processing tools use this data to provide snapshots of the collected data.
- Date/Time – The crux of the issue often involves a known relevant time period where both parties expect the data to exist. When data is processed, if an unusual amount of data exists outside of the time period or doesn't exist when it should, it provides the first indicia that something is amiss.
- Data relationships – Most processing tools provide textual analysis to see how the data is being grouped and its relationship to other words. This provides insight regarding whether the search term list needs to be revised. In addition, tools can also provide snapshots of who custodians talk to frequently which can aid in understanding your client's day to day operations.

At Trustpoint, our data is processed though Mindseye which allows us to get a deeper insight into the data yielding better results for you.

- Global Deduplication
- Email Threading
- DeNisting
- Concept Searching
- Near-Duplicate Identification
- Highlighting Data relationships

Reducing Noise – Using Mindseye to pare down the data allows you the ability to promote only the information that needs to be reviewed.

- Trustpoint can provide load files and managed review services in any review platform.

- Trustpoint can also provide customizable production processes to meet specific client needs for load file format, delivery type, stamping, metadata requirements and other parameters.