



Monday, April 30
2:30-4:00 pm

**2004 Getting the Information You Need:
eDiscovery, Conducting Interviews, Handling
Electronic Data & More**
Paralegal Track

Kay N. Pfeiffer
Senior Paralegal
The Dow Chemical Company

Andreae M. Selleck
Corporate Paralegal - Litigation Support
DuPont Legal

Sonya L. Sigler
General Counsel
Cataphora Inc.

Case Name

Via Phone ___

In Person ___

GENERAL INFORMATION

1. NAME/LOCATION/PHONE: (Business card):

2. CURRENT TITLE (HOW LONG IN POSITION):

3. SECRETARY/ASSISTANT'S NAME LOCATION/PHONE:

3. WHO WAS YOUR PREDECESSOR IN YOUR POSITION?

4. WHAT OTHER POSITIONS HAVE YOU HELD WITHIN THE COMPANY:

5. HAVE YOU ACQUIRED ANY ONE ELSE'S FILES, IF SO WHOSE:

6. DO YOU HAVE ANY FILES MAINTAINED IN ANY OTHER LOCATION (i.e. FILE ROOM, WAREHOUSE, OR OUTSIDE STORAGE): DOES YOUR SECRETARY MAINTAIN ANY FILES? IF SO, WHERE?

7. DO YOU WORK FROM A REMOTE LOCATION (HOME OFFICE)?

DOCUMENT INFORMATION

8. HAVE YOU SEEN A COPY OF THE HOLD ORDER FOR THE MATTER? (IF NO, SHOW THEM THE HOLD ORDER AND DO THEY HAVE ANY DOCUMENTS RELATING TO THIS LITIGATION)

10. WHAT WAS YOUR INVOLVEMENT IN THIS LITIGATION

11. WHAT TYPES OF DOCUMENTS HAVE BEEN GENERATED THAT WOULD BE SUBJECT TO THE HOLD ORDER? IF YES, PLEASE DESCRIBE -

Working Files
Correspondence
Memoranda
Reports
Handwritten Documents
Other

12. DO YOU HAVE ANY E-MAILS, DISKS OR OTHER ELECTRONIC FILES RELATING TO THE LITIGATION? (MAKE SURE YOU GET INFORMATION THAT YOUR "IT" PERSON NEEDS TO COPY E-MAIL ACCOUNTS – SERVER LOCATION, ACCOUNT NAME, ETC.)

(If answer is "yes", then ask the next questions)

13. WHAT E-MAIL SOFTWARE/APPLICATION DO YOU USE – COMPANY E-MAIL/ INTERNET E-MAIL (AOL, YAHOO)?

14. DO YOU ARCHIVE YOUR E-MAIL (if yes collect archived e-mail during interview)?

15. WHAT OTHER TYPES OF SOFTWARE DO YOU USE (i.e. Word, PowerPoint, Excel) LIST VERSION:

16. ARE YOUR DOCUMENTS ENCRYPTED OR DO YOU PASSWORD PROTECT FILES, IF YES WHAT TYPES OF FILES (PROVIDE PASSWORD):

17. WHERE DO YOU STORE/SAVE YOUR ELECTRONIC NON E-MAIL DOCUMENTS (i.e. My Documents folder, shared drive, Desktop etc.):

18. DO YOU USE SHARED TEAM ROOMS:

19. WHAT DATABASE APPLICATIONS DO YOU USE:

20. DO YOU STORE DATA ON ANY OTHER MEDIA (i.e. JAZ, FLOPPY, CD-ROM, FLASH/THUMB DRIVES OR REMOVABLE HARD DRIVES, PALM PILOTS,

PDAs, CALENDARS, BLACKBERRYS, MOBILE AND/OR CELL PHONES INCLUDING TEXTS AND DIGITAL PICTURES, DIGITAL CAMERAS):

21. DO YOU USE A CELL PHONE FOR WORK (If yes get cell provider's name):

22. HAVE YOU RECENTLY PURGED ANY FILES THAT MAY HAVE BEEN SUBJECT TO THE HOLD ORDER (If yes what types):

23. YOU ARE AWARE THAT SINCE THE HOLD ORDER HAS BEEN ISSUED YOU ARE RESPONSIBLE FOR MAINTAINING ANY AND ALL DOCUMENTS WHICH ARE SUBJECT TO THIS LITIGATION.

24. DO YOU KNOW OF ANYONE ELSE NOT LISTED ON THE DISTRIBUTION LIST OF THE HOLD ORDER WHO MAY HAVE FILES RELATING TO THIS LITIGATION:

25. NOTES/COMMENTS

Reminders:

Do not destroy any documents described in the collection memo you received prior to the interview
Please keep all new documents in a separate folder for future collection
Please remember that this meeting and the information you have given are privileged and confidential.
Please direct any questions concerning this interview to In-House Legal and/or law firm representatives.

Name of Interviewer: _____

Date of Interview: _____

The follow box numbers were assigned to this interviewee:

Box _____ to _____

EDD copied _____

of Files _____

***PRIVILEGED & CONFIDENTIAL
ATTORNEY WORK PRODUCT***

(Litigation Name) Box Tracking Log
(Plant/Site Name) Month/Year

Box # (One Custodian Per Box)	Location (Office Number or Location)	Custodian	Date Sent To Service Supplier	Date Returned to Site	Date Loaded to Database	Date Reviewed	Date QC	Date Produced

**PRIVILEGED AND CONFIDENTIAL
 ATTORNEY WORK PRODUCT**

In-House Document Review

December 6, 2006

Sonya L. Sigler

General Counsel

Agenda

- Rules Changes Summary
 - Paper
 - Electronically Stored Information
- *Assembling a Team*
- Discovery Process - How Is It Affected?
 - *Preservation*
 - *Review*
- Questions

FRCP 1

- **Rule 1. Scope and Purpose of Rules**
These rules... shall be construed and administered to secure the just, speedy, and inexpensive determination of every action.
- ... **two out of three ain't bad.**

FRCP Changes

- Rule 16 - Scheduling
- Rule 26 - Meet & Confer
- Rule 33 - Access to ESI
- Rule 34 - Sampling
- Rule 37 - Safe harbor?
- Rule 45 - Ordinarily Maintained, Not Reasonably Accessible

Follows Stages of Civil Litigation

- Commencement (FRCP 3 – 6)
- Pleadings (FRCP 7-15)
- **Pre-Trial / Scheduling Conference (FRCP 16)**
- **Discovery (FRCP 26-37)**
 - Of documents, including electronic data
 - Depositions
- Trial (FRCP 38-53)
- Post-trial (FRCP 54 – end)

Assembling an E-Discovery Team

- Include team members with diverse interests in the process, bringing specific priorities and responsibilities to their involvement with identifiable, yet changing roles
- Identify effective, flexible, experienced contenders
- Not a one-stop shop
- Understand that the team mix and/or roles may vary from project to project
- Include inside and outside expertise on legal and technical issues
- Ensure right mix of expertise for each project, each phase and avoid weak links

The Team

- Project Team v. Legal Team
- Essential Components
 - Corporate Stakeholder
 - Lead Attorney - outside counsel
 - Inside Counsel
 - IT Representative
 - Project Leader

The Players

Others - Corporate

- Records Management
- Compliance, Security, Risk Management
- HR
- Finance
- Accounting
- Operations

Others - Beyond Corporate

- Legal team - core group of associates, paralegals, in-house expertise
- Litigation support expertise - vendors, in-house, consultants, outside counsel
- Forensic/ data collection expertise
- Experts - testifying to process

Defining Roles

- Corporate Stakeholder
 - Risk management - representing fiscal interests
 - Analyze the spending and ROI
 - Identify gaps and/or weak links
 - Evaluate progress
 - Suggest process improvements
- **VITAL THAT LAWYER'S JUDGEMENT BE EXERCISED AT EACH STEP OF THE PROCESS**

Defining Roles

- Lead Attorney - Outside Counsel
 - Must know at least a little about every step of the process
 - Understand overall goals
 - With other corporate stakeholders
 - Analyze spending & ROI
 - Identify gaps and/or weak links
 - Evaluate process
 - Suggest process improvements
- **VITAL THAT LAWYER'S JUDGEMENT BE EXERCISED AT EACH STEP OF THE PROCESS**

Defining Roles

- IT Representative(s)
 - Thorough understanding of the corporate IT infrastructure
 - Ensure familiarity with actual practices, not just process and procedures
 - Manage internal IT resources essential to the process
 - Leverage to locate and identify data, but utilize those more qualified and experienced to gather, process, and review data

Defining Roles

- In-house Counsel
 - Time AND Focus
- Project Leader
 - Involved in every step of the process
 - Communication between legal team and project team go through project manager exclusively
 - Serves as a funnel for all decision making, process changes, etc.
 - Manages the project team NOT legal team

Communication and Coordination

- Have a plan (and a Plan B) and communicate it to ALL involved
 - Consider website (share point)
- Have kick-off meeting, face to face, if possible - team members see broader context and where they fit in
- Designate those with authority to make decisions - both *process* and *budget* related
- Utilize Project Leader as funnel

Document Every Step

- Scope and criteria - communicate modifications as project progresses
- Create an audit trail for changes
- Email group name - use for all Q&A or to document substance of oral conversations among team and subset of team members
- Utilize forms for tracking (chain of custody, status, disposition, etc.); make sure all team members have *access and use it*
- Designate shared repository - extranet, public folder, DMS, etc. - include all documentation

Rule 26 General Provisions Governing Discovery; Duty of Disclosure

- (a) **REQUIRED DISCLOSURES...**
a party **must**, without awaiting a discovery request, provide a copy of all documents, **electronically stored information**, and tangible things... in its control ***that it may use to support its claim or defense.***
- (b) **DISCOVERY SCOPE AND LIMITS. ... the scope of discovery is as follows:**
- (1) *In General.* Parties may obtain discovery regarding any matter, not privileged, ***...relevant to the claim or defense of any party***, including... any books, documents...

This is much more of a shield than 26(a)

“exponentially greater volume”

The discovery of electronically stored information raises markedly different issues from conventional discovery of paper records. Electronically stored information is characterized by exponentially greater volume than hard-copy documents. Commonly cited current examples of

From page 22 of the Commentary by the Rules Committee, Sept 2005

Electronic Data Collections Are...

- ◆ Volatile
 - ◆ Can change without your awareness
- ◆ Free-ranging
 - ◆ Not always located where you expect
- ◆ But, most importantly...
 - ◆ Prolific
 - ◆ Can duplicate rapidly

How Electronic Collections Are Different

- ◆ *Much larger data collections, but that's not all*
 - ◆ Meta data - Each item contains much data about itself, both hidden and visible
 - ◆ Meta-data is your friend
 - ◆ And there are *many* very small items
- ◆ Duplicates - Collections are often very largely redundant
- ◆ Mixture of informal and formal, personal and professional, highly useful and highly useless content, cryptic content
 - ◆ Email has replaced paper
 - ◆ Email has replaced phone calls

Stages of Discovery

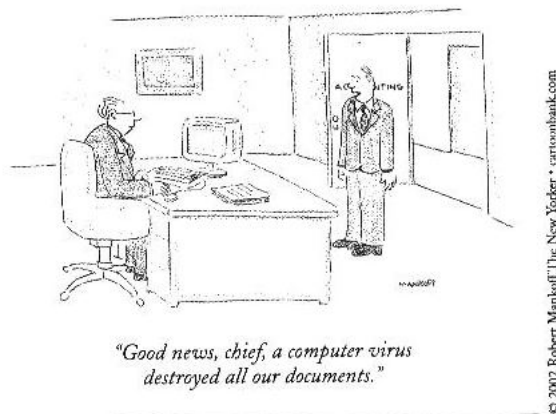
- **Preservation**
- Collection
- **Review**
- Production
- Investigation

Preservation

- Rule 26(f) Meet & Confer
 - Discovery Plan
 - Data Preservation
 - Scope - relevancy, custodians, date ranges
 - Accessible, Inaccessible
 - » File types, locations of data
 - Production of Data
 - Format, Other details
 - Privileged Information - Process
 - Upfront Discussion and Report
- Rule 16 - Scheduling Conference

38 FEDERAL RULES OF CIVIL PROCEDURE
 5 Discovery will be needed on the following subjects:
 6 _____ (brief description of subjects on which
 7 discovery will be needed) _____
 8 Disclosure or discovery of electronically stored
 9 information should be handled as follows: _____ (brief
 10 description of parties' proposals) _____
 11 The parties have agreed to an order regarding claims of
 12 privilege or of protection as trial-preparation material
 13 asserted after production, as follows: (brief description
 14 of provisions of proposed order)
 15 All discovery commenced in time to be completed by
 16 _____ (date) _____. [Discovery on _____ (issue for
 17 early discovery) _____ to be completed by
 18 _____ (date) _____.]

Obligation to Preserve



Rule 37 (f) - Safe Harbor?

Electronically stored information. Absent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide electronically stored information lost as a result of the routine, good faith operation of an electronic information system.

Spoliation that is **not** due to good faith:



Preservation

- Duty to Preserve
- Complying with Litigation Holds
 - Corporate Counsel Role Widened
 - Document Retention Requires More Attention
- Pull & Park Phenomena
 - Backup Tapes
 - FRCP - Safe Harbor - Rule 37
- Identify Not Reasonably Accessible Data

No more “I won’t ask if you don’t ask”

- **Electronic documents were discoverable under existing case law anyway, but....**
- **The new amendments remove any doubt**
- **... and all excuses.**
- **Now you do not dare “not ask.”**

Collection

- Practical
 - Cannot save the world of data - relevancy
 - Integrate with document management system, data archives, litigation hold process
- Closer relationship between Legal/IT
 - 30(b)6 depositions of IT
 - Where is the data
- Ordinarily maintained
- Reasonably Accessible

Not Reasonably Accessible Data

- **Rule 34. Production of Documents, Electronically Stored Information, and Things and Entry Upon Land for Inspection and Other Purposes.**

(a) SCOPE. Any party may serve on any other party a request (1) to produce and permit... to inspect, copy, test or sample, any designated documents or electronically stored information... stored in any medium.
- **Proactive Sampling v. Requesting Party Sampling**

Review

The Evolution of Electronic Discovery Processing

Print and Review 1995 AD	Print, Scan, Code, Review 1997	TIFF And Review Circa 1999	Simple Filtering 2000	Keyword Searching 2001	Analytical Defensible Reliable Reduction, then Review, 2002 + beyond
---	---	---	--------------------------------------	---------------------------------------	---

Sedona Conference

Search and Information Retrieval, Principle 1:

In litigation... where the volume of discoverable electronically stored information is large, it may not be feasible to perform human review of every document for responsiveness or privilege, and automated search and information retrieval methods and tools may be necessary and valuable.

***This isn't just a brainstorm
of words and phrases.***

Purpose of review

- Documents need to be selected
- Not all documents are responsive and/or relevant – estimated 85% are not
- Identify privileged and confidential documents
- Understand the issues of the case
- Locate smoking guns

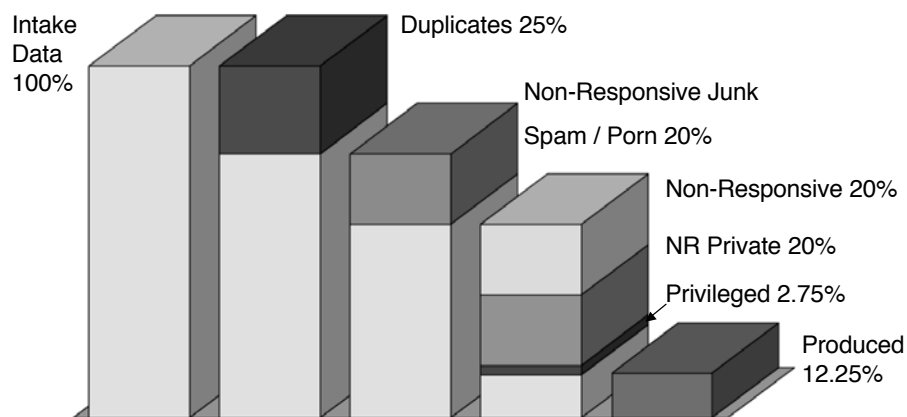
“E” in e-review

- Electronic documents are different from paper
- Many capabilities for electronic documents
 - E.g., searching, sorting, deduplication, near duplicate detection
- Take advantage of tools available
- Review in native format vs. imaged conversion (pdf, tiff, html, etc.)

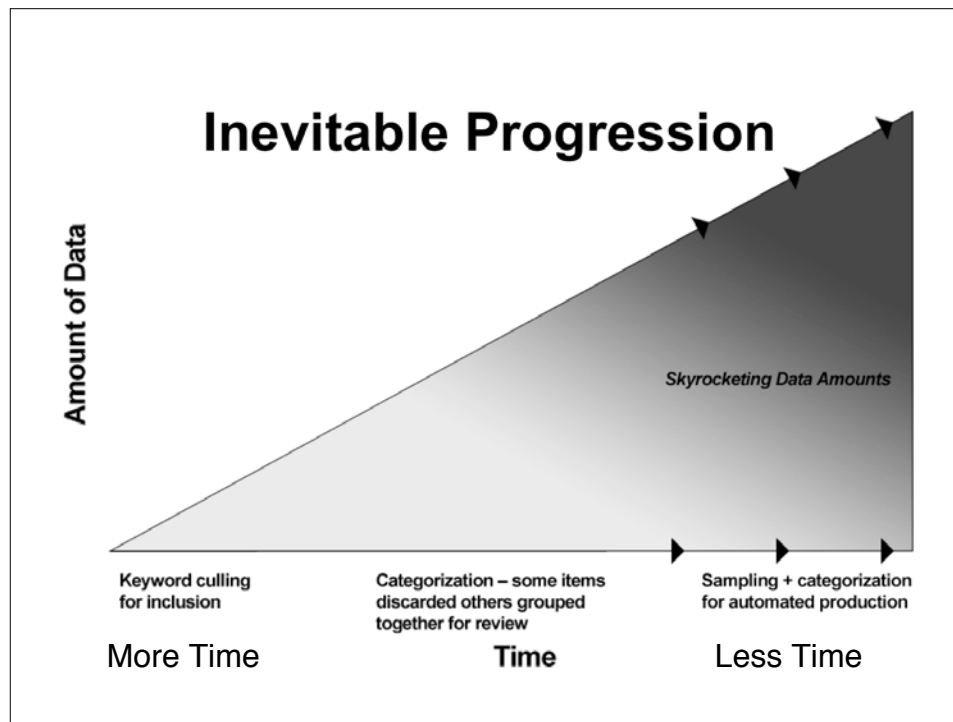
Review methodology

- Identify issues in the case
 - Formulate queries and approaches for finding responsive documents
 - Formulate relevancy and responsiveness guidelines
- Identify primary participants
- Select or triage documents for review

Getting to Reviewable Data

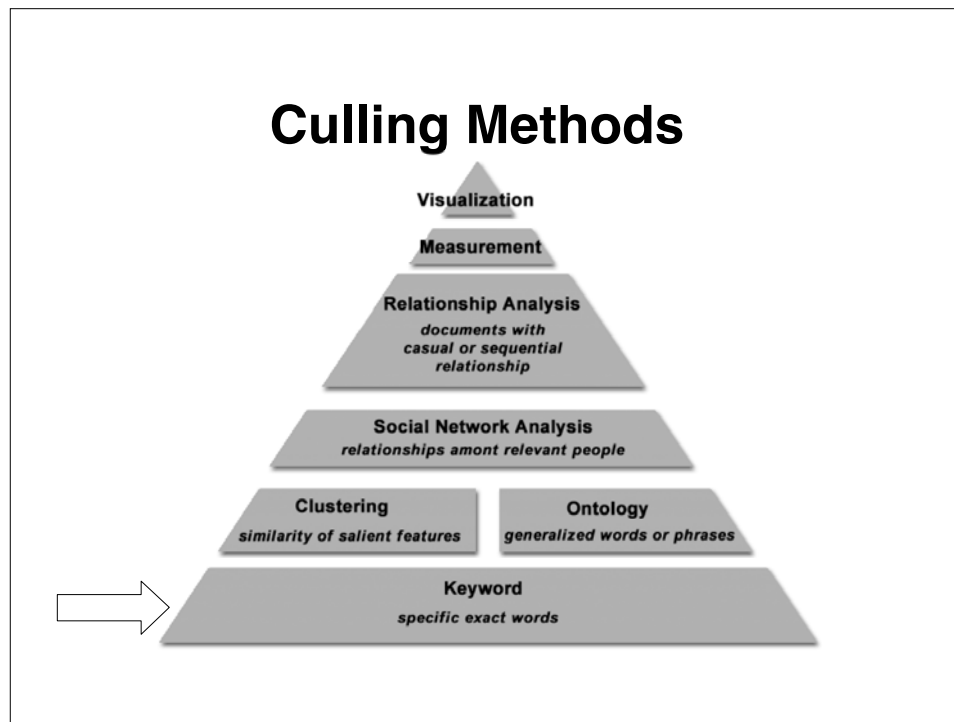


These figures vary based upon the data set received



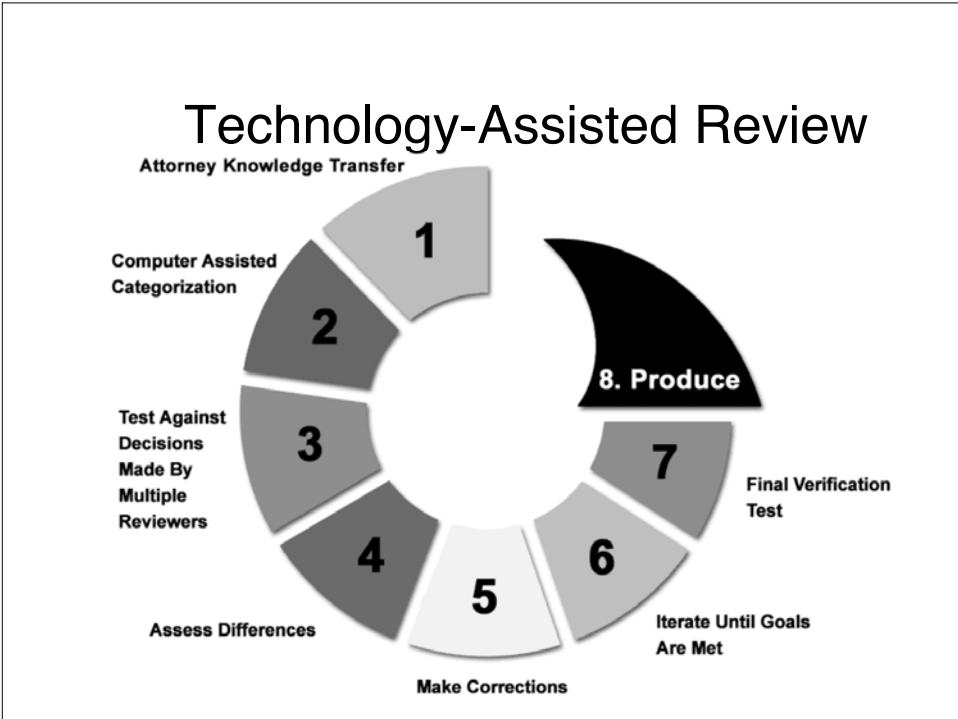
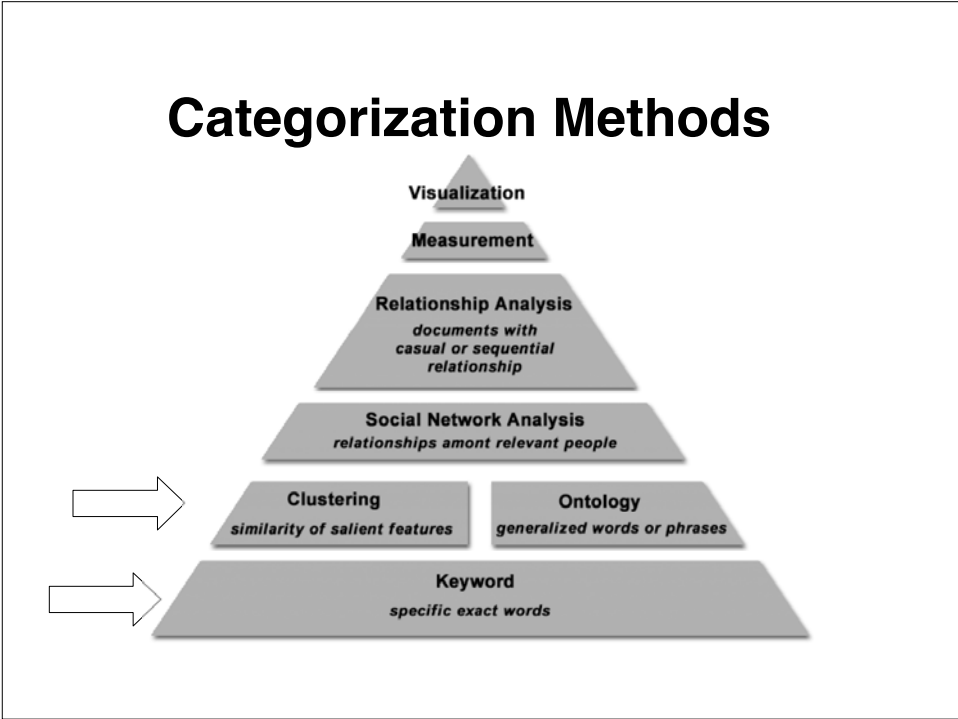
Review - Relevancy Assessment

- Beginning to End No Longer Tenable
 - 600 GB (10GB/60 custodians)
 - 100 reviewers almost 1 year to review it all
- Keyword searches, culling
- Categorization of Data
- Technology-Assisted Review
 - Categorization
 - Sampling



Categorization of Data for Review

- Categorize Entire Data Set
 - Spam/Porn/System Files
 - Personal/Private Data
 - Non-relevant Business Data
- Business Data
 - Relevancy Assessment
 - Privilege Review
- Keyword Analysis - Overlap, Holes



What can you learn from the
data beyond review?

Early Case Assessment

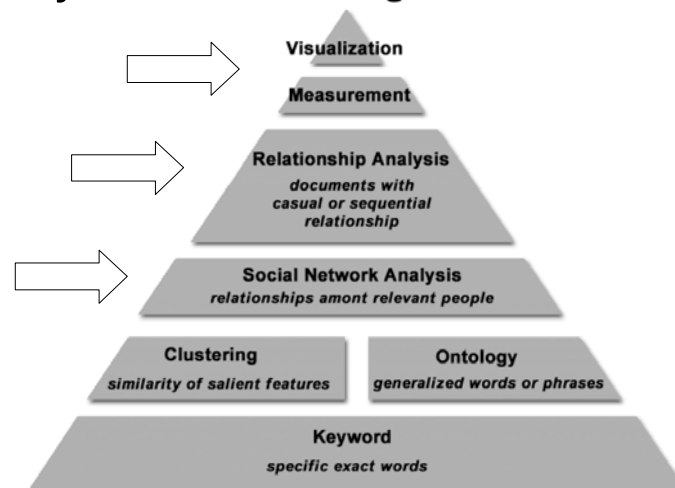
Analytics & Investigation - Making Sense of the Data

- Patterns/Variations
 - Counting (tally)
 - Analytics
 - Visualizations

Review tools

- What kinds of tools are available
 - Keyword search
 - Thesaurus
 - Taxonomy
 - Ontology
 - Clustering
 - Bayesian Classifiers
 - Concept-related tools
 - Latent Semantic Indexing
- How can they be used? Help to triage documents for review
- Automation? Use these tools to select documents for review or production

Analytics & Investigation Methods



Analytics In Report Form

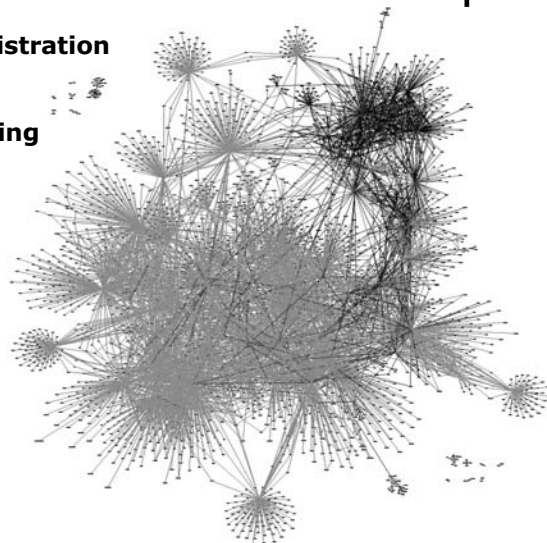
Name	Email	Company	Title	Other Addresses	Top 5 Closest Actors
Charles Berber	cb@jmc.com	JMC	CTO	cb23@aol.com	Vicki Chinn Vivian Thomas Kathleen Meyer Ted Johnson Frank Miller
Joe Williams	jw@jmc.com	JMC	CIO	n/a	Vivian Thomas Charles Berber Lucinda Williams Vicki Chinn Ted Johnson
Ted Johnson	tj@jmc.com	JMC	CEO	ted777@earthlink.net ted777@yahoo.com pingme04	Charles Berber Vivian Thomas Vicki Chinn Lucinda Williams Joe Williams
Vicki Chinn	vc@jmc.com	JMC	VP, Marketing	vicki2039@hotmail.com happie56	Vivian Thomas Ted Johnson Frank Miller Kathleen Meyer
Vivian Thomas	vt@jmc.com	JMC	Principal	n/a	Vicki Chinn Frank Miller Joe Williams Ted Johnson Frank Thomas

E-Mail Communications: The Entire Dataset – Up Front

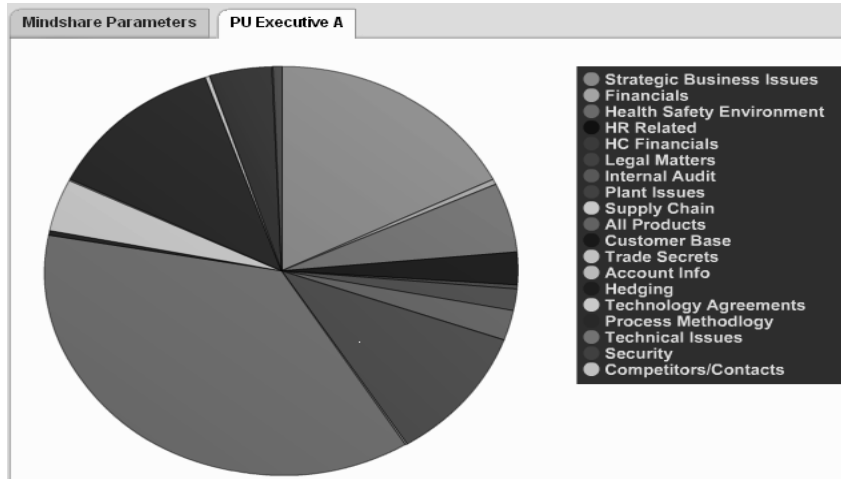
Green: Administration

Red: Legal

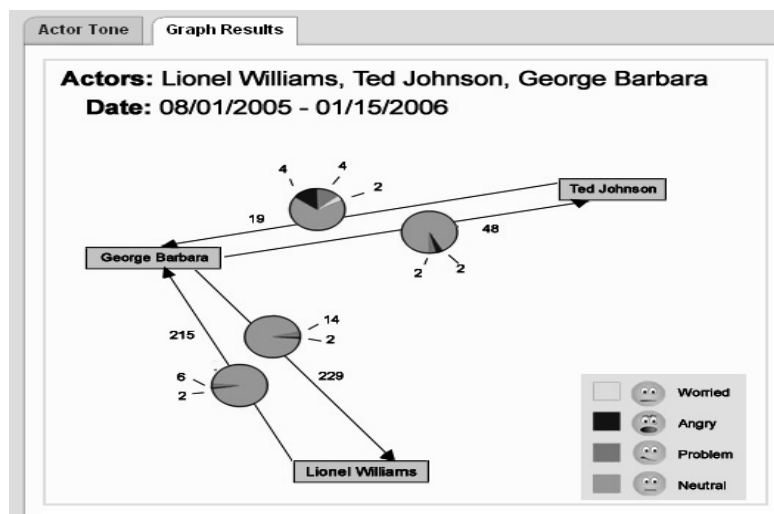
Blue: Accounting



Analytics - Mindshare



Analytics - Tone Detection



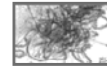
Analytics - Text Deletions



Amy Lawson



Juan Higuera



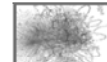
Steve Ahner



Frank James



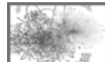
Neil Howard



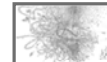
Ted Johnson



George Barbara



Sophia Teller



Vicki Chinn

Privilege

- Define the Process - Rule 26(b)(5)(B)
- State It Up Front
 - Quick Peek - turn over data quickly
 - Claw Back Procedures
- Follow the Procedure
 - Return
 - Sequester
 - Destroy
- Priv Log - how flexible is it?
 - Be Descriptive

Rule 26. General Provisions Governing Discovery; Duty of Disclosure

(b) Discovery Scope and Limits.

(5) Claims of Privilege or Protection of Trial Preparation Materials.

(B) Information produced. If information is produced in discovery that is subject to a claim of privilege... the party making the claim may notify any receiving party of the claim and its basis. After being notified, a party must promptly return, sequester, or destroy the specified information ... and may not use or disclose it until the claim is resolved.

This is the “claw-back”.

Review risks

- Failure to find responsive documents
- Failure to recognize responsive documents
- Failure to recognize privileged documents
- Inconsistent treatment of documents (e.g., duplicates)
- Failure to complete project in a timely manner

Evaluating the review process

- Accuracy of review depends on
 - Human factors (e.g., reviewer attention, knowledge, effort)
 - Query quality (if you don't ask the right questions, you cannot get the right answer)
 - Processing accuracy (e.g., miss email bodies or names on CC: list)
- Must evaluate all parts of the process
- Sampling techniques are available to assess quality of project

Possible **Production** Formats

- **Paper** (if the other side asks for it this way, be happy to oblige. It is the least useful format in which to receive a production.)
- **Paper-like (TIFF or PDF images)**
 - TIFF images without any searchable data at all are increasingly unacceptable.
- **Native Files**
- **Hosted “production” areas of the producing party’s web repository.**

FRCP 34 [amended, continued]

- **Rule 34. Production of Documents, Electronically Stored Information.**
 - (b) PROCEDURE. ... the request may specify the form in which the electronically stored information is to be produced....
 - ...[the responding party may object] to the requested form, stating the reasons, and the form it intends to use [instead].
 - (ii) if a request does not specify, the responding party must produce in the form it is *ordinarily maintained* or in an *electronically searchable form*.

FRCP Changes for Production

- Native
 - Meta-data and embedded data
 - Unique ID
- FRCP Default - Rule 34
 - Requesting Party Specifies or Parties Agree
 - Nothing Specified or agreed to
 - Ordinarily Maintained
 - Reasonably Usable
 - Not More than One Format
- Shared Database

Key Changes

- Introduces Term
 - “Electronically stored information” (ESI)
- Different requirements for paper and electronically stored information
 - Introduces concept of accessibility
 - Shift of burden to responding party for ESI
 - Inadvertent production of privileged material (“claw-back”) R.26(b)(5)(B)
 - Limiting sanctions for spoliation if producing party shows good faith efforts (“safe harbor”) R.37(f)
- Dramatic shift from prior practice
 - Legislates meeting and discussing plan for ESI

Key Concerns

- Uninformed bench
- Uninformed counsel
- Minimal precedent & guidance
 - Highly technical area with little clarity
 - Intangible subject matter
- Orders of magnitude increases in information
- How evolved is your data analysis and review technology? The other side’s?
 - May change depending on the matter

Drive early attention to e-discovery issues in scheduling conferences, orders, initial disclosures, discovery conferences and orders	16(b) 26(a) 26(f)
Establish a two-tiered approach to e-discovery based on accessibility of data: accessible ESI is handled as before; not reasonably accessible ESI is handled <i>differently</i> – a two tier system	26(b)(2)
Whether ESI is not reasonably accessible is a function of whether there is undue burden or cost in collecting the ESI	
Allow the responding party to not produce from sources <i>it identifies</i> as not reasonably accessible because of burden or cost	

In a motion to compel production of ESI the responding party has the burden of showing that the ESI is not reasonably accessible because of undue burden or cost; <i>this is different from paper</i>	26(b)(2)
Requesting party can still prevail with showing of good cause	
Court may specify conditions for any further discovery allowed: cost shifting, for example	
"A party's identification of sources of electronically stored information as not reasonably accessible does not relieve the party of its common-law or statutory duties to preserve evidence."	Duty to Preserve

Provide a means of addressing assertions of privilege after production	26(b)(5)
Allow interrogatories to be answered by production of ESI; <i>would mean granting access to systems or creating replicate systems</i>	33(d)
Distinguish electronically stored information (ESI) from documents; ESI includes any type of information that is stored electronically – <i>including metadata</i>	34(a)
Give requesting party the right to specify the form of production of ESI; <i>different than with paper productions</i>	34(b)

Provide, where not otherwise specified, that a party must produce ESI in a form in which it is <i>ordinarily maintained, or in a form that is reasonably useable</i> ; need not produce same ESI in more than one form	34(b)
Creates a safe harbor for routine good faith operations: “Absent exceptional circumstances, a court may not impose sanctions...for failing to provide ESI lost as a result of the routine, good faith operation of an ESI system”	37(f)
Conform subpoena practice to party discovery amendments; <i>places burden on the responding party to show the information not reasonably accessible because of undue burden or cost</i>	45

Duty to Preserve	Backlog of ESI waiting for discovery to begin which is subject to discovery in other matters (Pull & park)
	Fuzzy definition of scope of duty to preserve at start of matters
General	Consistency across matters and law firms used
	Need to know both physical and IT geographies of company
33(d)	Would require allowing access to home systems or creation of exact duplicate system for review work
26(b)(5)	Hollow privilege protection provision – may be used but not relied upon because not consistent with all state laws

34(b)	Form of production: <ul style="list-style-type: none"> ■ Integrity of native files ■ Native v. .tiff ■ Hardcopy v. electronic with same content ■ How to produce metadata
	Need to produce ESI in a form in which it is ordinarily maintained, or in a form that is reasonably useable
16(b) 26(a) 26(f)	Heightened expectations regarding case and discovery management, initial disclosures – need to know early on what your plan is
37(f)	The unsafe “safe harbor”

Informational Retrieval Glossary



Boolean

Boolean search is a common technique which allows terms to be logically connected using AND, OR and NOT to refine a query. For example: "(mercury OR saturn) AND NOT automobile" would return all documents containing either the word mercury or the word saturn and not containing the word automobile.

Bayesian Classification

In general, bayesian classification is based on the statistical probability of a class and the features associated with that class. This type of classification utilizes a training set composed of classes that have correctly assigned features. Once the probabilities of the training set features and classes have been stored, new data is compared against the training set. During this comparison of the "learned" classification of the training set with the new data, the new data's features are calculated and the new data are assigned classes whose probability of matching the training set's classes and features is highest.

Categorization

In general, categorization is the grouping of objects, people or ideas on the basis of some kind of "similarity". As applied to electronic discovery, it describes the grouping of documents according to some desired criteria. Categorization may be by topic or by legal criteria. Categorizing documents about specific products, or documents that relate to sales in a given country are examples of categorizing by topic. Legal criteria might include categorizing responsive, non-responsive, and privileged documents.

Clustering

Clustering is the grouping of information by some category or statistical similarity. This is done through various grammatical, semantic, and even punctuation algorithms designed by combination to detect topics rather than just individual keywords. Statistical clustering can be done by counting words and their frequency, then grouping those documents with similar statistics together in a cluster. When files are determined to be about the same or similar topics, they are clustered together, and usually displayed in some kind of graphical relationship that facilitates reviewing similar documents together.

Concept

Concept search attempts to find documents that address some concept that a user is interested in. To do so, it goes beyond *keyword* search for documents that contain a specified word or phrase, and tries to find other documents that address the underlying concept. For example, a concept search for "fiber" might return documents that refer to the concept of fiber using alternative terms such as cloth, material, cotton etc.

Keyword

Keyword search looks for documents that contain a specific word or phrase. Keyword searches may be further refined by using *Boolean* operators.

Latent Semantic Indexing

Latent Semantic Indexing has involves extracting multiple concepts from the data collections through a statistical semantic analysis of each file. The theory is that unstructured files comprise *latent* concepts that are not readily recognized and remain hidden until a more precise lexicon is developed out of the whole collection. These concepts then form a dictionary (lexicon) for the collection that can be weighted for both frequency of occurrence and relevance. At that point each file in the collection is compared to the concepts list, and it is assigned a *fingerprint* (or value) that uniquely defines the file according to those criteria. Searches can then be conducted by requesting files that are statistically similar, i.e. that have similar fingerprints, under the presumption they will be not just similar but conceptually related as well.

Informational Retrieval Glossary



Linguistic Techniques

Search or categorization techniques that are based on analysis of language features of documents, in contrast with *statistical* techniques. *Ontologies* are an example of a linguistic technique.

Neural Network

A neural network is a computer program whose operation is loosely inspired by the way a human or animal brain works (though the neural network is much, much simpler). A neural network can be "trained" by giving it sample inputs and the correct outputs associated with these. The network can analyze the difference between the answers it is generating and the "correct" answers. It can then automatically adjust its internal workings, until its answers on the training set adequately match the given outputs. The idea is that you can now feed it new inputs (the answers to which are unknown) and it should now be able to provide the correct outputs for these. For purposes of electronic discovery, the inputs might be information about documents and the outputs a *categorization* of those documents.

Ontology

An ontology is an arrangement of words, phrases and search terms under a *concept*. Here is a simple example:

AIRCRAFT CONCEPT

- Boeing
- 747
- Cessna
- Glider

By reading documents and establishing whether they contain any of the four terms listed under the AIRCRAFT CONCEPT, we can determine whether any of the documents discuss the concept of aircraft. This process can be automated, so that a computer does the work. If the computer finds a document that contains one or more of the four terms, it concludes that the document is (at least partially) about aircraft. The document might also discuss other concepts, but a reference to the concept of aircraft is clearly present in the document.

Relevant

"Having some reasonable connection with, and in regard to evidence in trial, having some value or tendency to prove a matter of fact significant to the case." Finding relevant documents more effectively than our competitors is a key value that we provide.

Related

Documents that may be related to each other in a number of ways, such as addressing the same issue, or being created by a significant *actor* at a crucial time. Most electronic discovery software can find only documents which are similar to each other, or which belong to the same *email thread*. Cataphora *Discussions* are the first technology to pull together related documents, not just similar ones.

Statistical

Categorization may be based on statistical analysis of the similarity of documents. For this, a document is mathematically represented by a set of features such as the occurrence of words, or their proximity to other words in the documents. Different weights (levels of importance) may be assigned to the various features. Documents are then deemed to be similar (and therefore belong to the same category), based on the degree to which their features resemble each other.



Informational Retrieval Glossary

Taxonomy

Taxonomy is the practice and science of classification. Taxonomies, which are composed of *taxonomic units* known as *taxa* (singular *taxon*), are frequently hierarchical in structure, commonly displaying parent-child relationships.

Vector Space Modeling

Vector Space Modeling (VSM) is a concept that first came into favor in the early 1970s and it has provided some additional guidance in automated document review even to this day. It is based on building vectors that describe the relationships between each search query and each file in the collection. Each vector, by its magnitude and direction then maps to other files that are closest to it in relation to the same *feature* as emphasized by the search query. Each file thus becomes a compilation of *features* that place it in a multi-dimensional construct. That construct can be realized in a graphical display depicting all the relationships as vector lines between and among separate files.



cataphora

Data Collection, File Types, and Costs

Data Collection		File Types		Costs	
Checklist(s)					
<p>A. Places where to look for data collections:</p> <ul style="list-style-type: none"> - Shared Servers/ or Network - Desktop(s) - Laptop(s)/ Notebooks - PDA (Personal Digital Assistant)/ handheld devices - Notebooks - Backup Tapes - Storage <p>Locations:</p> <ul style="list-style-type: none"> ~ file shares, ~email devices, ~archival tapes, ~hosted emails, ~ attachments 	<p>B. Kinds of data to collect /or not to collect (file types):</p> <p>Archive Container Files</p> <p>Compressed archive(s). Only content of the compressed archive is extracted for further processing.</p>	<p>Supported</p> <ul style="list-style-type: none"> - Archive.iso: An optical disk image or a Universal Disk Format (UDF) image. It contains all the data files and the filesystem metadata in a single file. - Archive.ace - Archive.arj - Archive.bzz - Archive.jar - Archive.lzh - Archive.cab - Archive.tar <p>This file type can contain only one file or many files in multiple directories.</p> <ul style="list-style-type: none"> - Archive.gz: Generated using the GNU zip program. - Archive.iso: An optical disk image or a Universal Disk Format (UDF) image. It contains all the data files and the filesystem metadata in a single file. - Archive.rar: A compressed archive. RAR is a compressed file like ZIP, but uses a different encoding format. - Archive.uue: Compressed ASCII file archive created by UUENCODE (uuexe515.exe) 	<p>C. Single vs. Multiple Vendors</p> <p>General Fees</p> <p>Training Costs</p> <p>Data Loading and Processing Costs</p>	<ul style="list-style-type: none"> - Using different vendors for various aspects of the projects might decrease the cost. - More efficient to use vendors that can integrate with in-house applications to reduce loading and hosting costs. - Set up Costs <ul style="list-style-type: none"> • Fixed fee • Hourly (consider a cap on fees) - User license fees (month vs. months) - License subscription fees - Live Training <ul style="list-style-type: none"> • Day rates vs. hourly rates • Travel costs • Vendor's waiver for some of training sessions - Web based training <ul style="list-style-type: none"> • Session fee vs. hourly fee • Free web-based training by vendor - Identify what terminology vendor is using and what activities are included in price. - Are services bundled together for one fee, or assessed separately? - Loading Costs <ul style="list-style-type: none"> • Native • TIFF/PDF 	



cataphora

Data Collection, File Types, and Costs

Data Collection		File Types		Costs		
Checklist(s)						
		<p>Compressed archive(s). This file type can contain only one file or many files in multiple directories. Only content of the compressed archive is extracted for further processing.</p>	<p>- Archive.z: UNIX Compressed Archive File</p> <p>- Archive.zip: Zip files compress data and therefore save time and space and make downloading software and transferring e-mail attachments faster.</p> <p>Not Supported - Archive: <i>This file type can contain one file or many files in multiple directories. Unknown format.</i></p> <p>- Archive.exe.zip: <i>A self extractable zip archive or any executable that may contain zip archive.</i></p> <p>- Archive.ish: <i>Compressed archive. This is not very commonly used.</i></p>	<p>Data Loading and Processing Costs</p> <p>Hosting Costs</p> <p>Best Practice to Reduce Hosting Cost</p> <p>Hosting Shared Databases</p> <p>Costs of Archiving Data</p> <p>Costs of Production</p>	<p>- Costs Assessed - Other processing/loading issues - Common for volume discounts. Negotiate reduced rates as volume increases.</p> <p>- Data hosting charged per volume vs. monthly basis. - Negotiate</p> <p>- Reduce volume of data as much as possible - Review sources of data prior to loading to exclude duplicate data - Use vendor to hold unique copy of document</p> <p>- Ability of vendor to archive data for a dormant case for a reduced hosting fee - Cost for re-activating database</p> <p>- Dependent on format - Reduced rate to reproduce earlier production - Additional Costs - Paper production Costs (per page)</p>	
		<p>Not Supported</p> <p>Email Container Files</p> <p>Only content of the NOTES file is extracted for further processing. MBOX files fully processed.</p>	<p>Supported</p> <p>LotusNotes: A Lotus Notes database file. This type contains email folders, calendar items, tasks, custom forms, contact information, or other files that are stored in the database. Lotus Notes and Domino are a fundamental component of IBM Workplace.</p>			



cataphora

Data Collection, File Types, and Costs

Data Collection		File Types		Costs	
Checklist(s)					
		<p>Only content of the MBOX/MSG file is extracted for further processing. Fully processed.</p> <p>Converted to RFC-822 compliant MBOX format.</p> <p><i>Email archive</i></p> <p><i>This type includes all email folders, calendar items, tasks or contact information that Outlook Express collects.</i></p> <p>Not Supported</p>	<p>- Mbox: RFC-822 compliant format that may contain one or more email messages, calendar items, etc. This file type is a standardized email format.</p> <p>- Outlook.msg: MSG files are outlook messages saved as files, or encapsulated email messages.</p> <p>- Outlook.pst: This type includes all email folders, calendar items, tasks or contact information that Outlook 97, 2K, XP, or 2003 collects</p> <p>Not Supported</p> <p>- Eudora</p> <p>- OE(Outlook Express)</p> <p>- OE.dbx</p> <p>- OE.idx</p> <p>- OE.mbx</p> <p>- Outlook.pab: <i>PAB is the file extension for the Personal Address Book associated with Microsoft e-mail programs.</i></p> <p>Outlook.ost: <i>OST is the file extension for offline storage files associated with Microsoft Exchange/Outlook.</i></p>	<p>Professional Services Fees</p> <p>Cost Savings in Professional Service Fees</p> <p>Termination Costs</p> <p>Service Level Agreements</p>	<p>- Charged per hour.</p> <p>- Unpredictable costs.</p> <p>- Certain services can be performed by other persons (in-house counsel, outside counsel, litigation support personnel or consultant.</p> <p>- Sliding fee/level of experience</p> <p>- Blocks of hours for reduced fee</p> <p>- Build in level of support, or hours of support, into loading charges.</p> <p>- Costs for closing down database or transferring the data and work product to another vendor</p> <p>- Service Level Agreements provide for certain level of service guaranteed by financial penalties.</p>



cataphora

Data Collection, File Types, and Costs

Data Collection		File Types		Costs	
Checklist(s)					
		<p>Text Extractable Files</p> <p>Fully Processed</p> <p>Individual desktop application files. Files that are identified as part of Microsoft Word family.</p>	<p>Processed</p> <p>- ASCII: A text file with pure ASCII text and no formatting.</p> <p>- Code.c: C/C++ Program File</p> <p>- Code.java: Java Source Code</p> <p>- Code.js: JavaScript Source Code, This extension can be captured by a number of applications including: Windows Script Host, Dreamweaver MX, Notepad, Netscape Navigator, PavScrip, UltraEdit. The files are generally text files.</p> <p>- Code.sql: Structured Query Language Data</p> <p>- Concordance: Single page tiff formats bounded by summation meta file documents.</p> <p>- Email</p> <p>- HTML</p> <p>- MSOffice.doc</p> <p>- MSOffice.mpp</p> <p>- MSOffice.ppt</p> <p>- MSOffice.xls</p> <p>Fully processed</p>		



cataphora

Data Collection, File Types, and Costs

Data Collection		File Types		Costs	
Checklist(s)					
			<ul style="list-style-type: none"> - PDF - RTF - VSD - WPD - XML <p>Fully processed</p>		
		<p>Image Files</p> <p>OCR is available or delivered OCR text can be imported. If authorized, gathered text is indexed.</p>	<p>No action is performed</p> <ul style="list-style-type: none"> - Image.bmp - Image.gif - Image.jpg - Image.pcx - Image.png - Image.tif <p>A Tagged Image Format File (tiff). TIFF images are a common format when converting electronic/paper materials into a standardized storage or review mechanism.</p>		
		<p>Audio Files</p> <p><i>Not Supported</i></p>	<p>Not Supported</p> <ul style="list-style-type: none"> - Audio - Audio.aac - Audio.aif - Audio.mid - Audio.midi - Audio.mp3 - Audio.mpa - Audio.ra - Audio.ram - Audio.wav - Audio.wma 		
		<p><i>Not Supported</i></p>	<ul style="list-style-type: none"> - Video - Video.asf - Video.asx - Video.avi - Video.mov - Video.mp4 - Video.mpg - Video.qt - Video.rm 		



cataphora

Data Collection, File Types, and Costs

Data Collection		File Types		Costs	
Checklist(s)					
		Others	- Video.swf - Video.wmv No action is performed - Exe: A program file. Virtually all programs that run under Windows or DOS are in the .EXE format. - Known_os		
		<i>Not Supported</i>	Not Supported - XSL - MSAccess		

Freeze Letter Information for: *State Product or Matter***Attorney Contact:****Paralegal Contact:****Date Created:****Applies to:** *State entity that is affected by the freeze.***Freeze Start Date:** *The date that the freeze period begins***Freeze End Date:** *The date that the freeze period ends.***Product or Activity at Issue:**

A listing of the product(s) and/or activity(ies) giving rise to the freeze letter (antitrust investigation regarding product X, product liability action regarding product Y) that will alert the reader to the focus of the freeze.

Statement of Freeze:

Narrative statement of the scope and basis of the freeze, providing appropriate detail so the reader can understand the scope of the freeze in conjunction with the detailed listing of record titles.

Definition of Document *This is part of the template. It cannot be altered.*

The term “document” is very broadly defined. It includes all records or other materials in one’s possession or control, regardless of form, format or media. It includes the original and any non-identical copy. It includes all versions, not just the most recent or final. It includes anything that is written or otherwise recorded, such as books, pamphlets, periodicals, bulletins, calendars, diaries, minutes, memoranda, post-it notes, reports, studies, instructions, working papers, lab books, letters, meeting reports, contracts, graphs, charts, photographs and other images, microfilm, microfiche, x-rays, medical records, brochures, pamphlets, handwritten notes, working drafts, voicemails, videotapes, dictation tapes, electronic files and e-mails. It includes electronic documents, files, emails, instant messages and other electronic data compilations regardless of storage location or media, including information located on or in hard drives, network or system drives, file shares, personal digital assistants, cell phones, removable storage devices such as CDs, DVDs, USB cards, ZIP drives, and the like.

Collection and Records Handling *This is part of the template. It cannot be altered.*

Any actual collection of documents will be preceded by a collection letter setting forth the timing and details of such an effort. You do not need to do anything with responsive documents until you receive a collection letter, unless you find you have records with expired retention dates that are subject to this freeze letter. In such a case, please promptly inform one of the contact persons indicated. One of them will get in touch with you to make arrangements to collect the materials.

Expired records that are frozen are not to be retained in their home department.

You will be notified when this document freeze is no longer in effect.

Sample

Attorney - Client Communication

To: Distribution List

DOCUMENT REVIEW AND COLLECTION

[Insert Case Name or Project Name]

We have been ordered by the Court in the above-referenced case to provide documents relating to _____, _____ and _____ [Insert names of Products or other subjects within scope of request] (the "Collection Scope").

Consequently, we need to collect from you all paper documents relating to the Collection Scope mentioned above that have either been created by you or have come into your possession. Examples of relevant types of documents are set forth in Attachment A. Please search your files for any responsive documents such as those outlined in Attachment A. It would greatly assist us in our collection efforts if you would conduct a preliminary review for any and all responsive documents. Doing so will allow you to isolate extraneous materials and shorten our time in your area.

Please note that you will separately receive instruction for the collection of electronic documents and email.

A Legal Team will be visiting your office the week of _____, 200_, to discuss the case, review and image the responsive material and to answer any of your questions. We will call to set up a specific time that is convenient for you. *Could you please take a moment to reply to this message and fill out the questionnaire set forth on, Attachment B below by _____, 200_ [Insert date prior to office visit]. This information will assist us in drafting an argument to the Court on seeking to narrow the scope of this request and relieve some of the burden associated with this request. In addition, we intend to use this information to plan our staffing and the amount of time that we will need to spend in your area.*

When collecting documents, please remember to include hand-written notes, phonemail messages, e-mail messages, computer-stored documentation, day planners, photographic material, maps, etc. Also, if you believe you or your department may have relevant documents at the Records Center, please let us know. We will go to the Records Center and do the searching for you.

We recognize this activity imposes a burden on you, but it is both imperative and beneficial to for us to conduct this review and collection of files. Should you have any questions please contact either me or [Name], a Paralegal, at (____) _____. Please note all written communications directed to the Legal Department should be marked as "Attorney-Client Communication." Also, I would appreciate it if you would take a moment to review the distribution list attached to this request. If you think there is someone with relevant information who is not included on this list, please let us know and we will contact that individual.

Thank you for your cooperation with this collection. If you have any questions, please contact us.

Sincerely,

_____ [Insert name of case manager or discovery manager]

**

Attachment A
Categories of Relevant Documents

[INSERT RELEVANT DOCUMENT CATEGORIES TAILORED TO ADDRESS SCOPE OF COLLECTION. BELOW IS A SAMPLE ONLY]

**

[Attachment B](#)
[Questionnaire](#)

Please tell us what type of documents and files we will find in your office.

Please estimate the volume of responsive documents that are located in your office, i.e. number of file folders, drawers, shelves.

Please estimate the amount of time we would need to spend in your office to read the documents to determine whether they are responsive to this request.

Please estimate the volume of responsive documents that are located in your area (excluding your office). This would include documents located in Operating Discipline Rooms, Libraries and Common Files.

Please estimate the amount of time we would need to spend to read the document located in your area (excluding your office) to determine whether they are responsive to this request.

[Distribution List](#)

Procedure for Document Collection and Return

Code for Flag Colors

Green (go)	Responsive
Red (stop)	Not Responsive
Pink	Privilege
Blue	Oversized

Attorney Interviews Document Custodian

- Explain Case and Document Request
- Discuss electronic collection and location of files
- Discuss sources of electronic data (cd's; pda; treo...)
- We will use outcards and dictation to assure that documents are returned to their original location.
- Make sure to ask questions about documents that may be at Record Center
- Discuss sending docs to Utah
 - Two day turnaround
 - We will come back and return them to their files
- Discusses location of documents with Custodian

Attorney Reviews Documents

- Flag outside of file drawers
 - If it contains relevant documents – green
 - If no relevant documents – red
 - Privilege documents - pink
- Flag outside of file folders
 - If it contains relevant documents – green
 - If no relevant documents – red
 - Privilege documents - pink
- Flag documents inside of file folders
 - Use paper clips, binder clips to clearly separate documents
- Blue flags are used for oversize documents – maps, posters on the wall, presentation exhibits. Any large documents that would require special handling for processing. Larger than a copier could handle.

Paralegal Processes Documents

- Locate file drawers with green tags
- Locate file folders with green tags
 - Dictate the location of file drawer, name on file folder and title of relevant document
 - With the use of outcard, remove document and place in Paige box
- Note on Paralegal chart any oversized documents that we will have to come back for.
- Fill out “pink” sheet with number assigned for custodian, volume of documents, attorney reviewing docs.
- Pack in box, fill out FedEx form, take to mailroom and send to Utah for processing.
 - If custodian will not allow original documents to go, copy and send copies to Utah
 - Strongly encourage custodian to allow sending to Utah.
- Receive documents from Utah
- With the use of outcards and dictation, returns documents to file.
- Remove all tape flags from office.
- Note on Paralegal Chart that office is complete.



Andreae M. Sellek

- Setting up the Team
- Preservation
- Collection & Interview Process
- Selecting a vendor

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Financial Firms Are Struggling to Manage, Archive and Monitor the Exponentially Growing Volume of Corporate E-mail



By Paul McDougall and Elena Malykhina, InformationWeek
Wall Street & Technology
Sep 21, 2006

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Trigger

- Complaint
- Investigation
- Notice of Lawsuit
- Merger or Acquisition

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Team Members – Company

- Client/Business
- In-House Attorney(s)
- Paralegal(s)
- Litigation Support
- IT Personnel
- Other Employees
- Other Paralegals

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Company Policies/Guidelines

- Records Retention
- Records Hold
- Backup Tapes – E-mail/E-Documents
- Outside Counsel
- Information Security Organization/ IT
- Collection, Processing and Production of Documents

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Team Members – Law Firm

- Outside Counsel
 - Attorney(s)
 - Paralegal(s)
 - Litigation Support

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Team Members - Vendor

- Vendor
 - Preferred Vendors
 - Project Manager

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Team Members - Vendor

- Meeting with Vendor
 - Make Sure that Vendor Knows How to Process ESI
 - Make Sure that is Vendor Handling ESI so Metadata is Not Getting Changed
 - How is Vendor Handling Exceptions
 - How is ESI Stored
 - Ask for Budget Estimates

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Selection of a Vendor

- Requirements
 - Processing
 - Database Functionality
 - User Interface
 - Searching Capabilities
 - On-Line Review
 - Productions
 - Reports/Exports
- Demo Software

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Selection of a Vendor

- Interview Vendor's Other Clients
 - Name of Interviewee, Company/ Position
 - Are they an Active User of the Software
 - How is Database Technology Being Used
 - ESI or Paper
 - How Long have they Been Using this Software and Number of Cases/ Volume of Documents/ Pages.
 - What Functionality is Used Most

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Selection of a Vendor

- Have Deadlines been Met
- Project Management
- Pros
- Cons
- If there was one Thing you Could Change What Would it be
- Pilot Software

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Team Members - Contractors

- Team Lead
- Collection and Review Team

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Team Members

Once you have Your Team Members have a Contact List Created and Update it as Necessary.



Preservation

- Records Hold Order
 - Who Drafts the Records Hold Order
 - How is it Communicated
 - How Often is a Reminder Sent Out
- Revise Records Hold Order as Necessary



Records Hold Order

- Business/JV/Subsidiary
- Product(s)
- Region/Country
- Case Name and Matter Number
- Date Parameters
- Summary and Scope of Matter
- Description of Records Subject to Hold Order

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Penalties for Failure to Preserve

- Monetary Sanctions
- Adverse Inference Charge
- Preclusion of Evidence
- Default Judgment

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Interviews

- Identify Custodians/ Witnesses
 - Keep List of Custodians
- In Person and/or Telephone Interviews
- Copy of the Hold Order
- Interview Form

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Interview Form

- Contact Information
- Does their Assistant/Secretary Maintain Any Files
- What is Their Involvement in this Case
- What Types of Documents do They Have
- What Types of E-Mail Software/ Applications They Use

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Interview Form

- Archive E-Mail
- Instant Messaging
- Types of Software (i.e. Word, PowerPoint, excel)
- Database Applications

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Interview Form

- Password Protect/Encrypted Files, if yes
What Types of Files (Get Password)
- Store/Save Electronic Documents (i.e. My Documents Folder, Shared Drive, Desktop, etc.)

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Interview Form

- Other Stored data and media
 - Floppy Discs
 - CD-Rom
 - Flash/Thumb Drives or Removable Hard Drives,
 - Blackberries, Palm Pilots, Calendars PDAs,
 - Cell Phones (Text Messaging and Camera)
 - Digital Cameras
- Voice Mail

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Classify Custodians

- Add Any New Custodians to the List of Custodians that were Identified During the Interviews
- Before You Collect or Process **any** ESI Classify Your Custodians
 - Key
 - Relevant
 - Possibly Relevant
 - Not relevant

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Discuss Processing/ Production Strategy

- Company's Views on how to Process ESI
- Understand How ESI is Collected and Processed
- Volume of ESI to be Collected and Processed
- Value of the Case
- Form of Production

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Collection Plan

- Plan Early, if possible
- Communicate
- Consistent
- Defensible – Not Perfect
- Focused Collection
 - Training
- No Spoliation
- Document Decisions
- Proactive

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Collection Criteria

- Complaint
- Records Hold Order
- Request for Production
- Parameters/Scope
- Objections
- Deadlines
- Single Case or Serial Litigation

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Collection Plan

- Identified Potential Document Custodians and Locations,
- Interview Custodians
- **Don't Print or Forward Any Custodian's E-mails or Electronic Documents**
- Identify Where E-mails and ESI are located
 - Is there a map of where company has ESI
 - Are there any E-mail janitors

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Collection and Processing Plan

- Forensics
- Back Ups
- What is the Volume
 - Sampling
- Who is Doing the Collecting
 - Custodians
 - Collection Teams
- Are There Any Restrictions

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Collection Plan

- Track Hard Copy Documents and ESI
 - Chain of Custody
- Get Information you Need to Provide to IT
- Identify Unusual Forms of ESI
- Copy ESI
- Keep Pristine Copy of ESI

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Site Collection

- Obtain a Contact Person at the Site to Assist you
 - Gain Access to Site
 - Schedule Interviews
 - Assist with Supplies
 - On-site Shipping Information, etc.
- External Hard Drive for all Electronic Documents
 - Training

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Site Collection

- Inform Vendor of Your Plans so they can Coordinate
- Track What Leaves the Site and When it is Returned
- Plan to put Hard Copy Documents Back

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Sonya L. Sigler

- Changes in Federal Rules
 - How Does it Affect Discovery
- Processing
 - All This Data...
 - What to Do?
- Budgets
 - Where Does All that Money Go?

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



FRCP 1

- **Rule 1. Scope and Purpose of Rules**

These rules... shall be construed and administered to secure the just, speedy, and inexpensive determination of every action.
- ... **two out of three ain't bad.**

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



FRCP Changes

- Rule 16 - Scheduling
- Rule 26 - Meet & Confer
- Rule 33 - Access to ESI
- Rule 34 - Sampling
- Rule 37 - Safe harbor?
- Rule 45 - Ordinarily Maintained, Not Reasonably Accessible

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Follows Stages of Civil Litigation

- Commencement (FRCP 3 – 6)
- Pleadings (FRCP 7-15)
- **Pre-Trial / Scheduling Conference (FRCP 16)**
- **Discovery (FRCP 26-37)**
 - **Of documents, including electronic data**
 - **Depositions**
- Trial (FRCP 38-53)
- Post-trial (FRCP 54 – end)

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Rule 26 General Provisions Governing Discovery; Duty of Disclosure

- (a) **REQUIRED DISCLOSURES...**
a party **must**, without awaiting a discovery request, provide a copy of all documents, **electronically stored information**, and tangible things... in its control *that it may use to support its claim or defense*.
- (b) **DISCOVERY SCOPE AND LIMITS. ... the scope of discovery is as follows:**
- (1) *In General*. Parties may obtain discovery regarding any matter, not privileged,...*relevant to the claim or defense of any party*, including... any books, documents...

This is much more of a shield than 26(a)

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Preservation

- Rule 26(f) Meet & Confer
 - Discovery Plan
 - Data Preservation
 - Scope - relevancy, custodians, date ranges
 - Accessible, Inaccessible
 - » File types, locations of data
 - Production of Data
 - Format, Other details
 - Privileged Information - Process
 - Upfront Discussion and Report
- Rule 16 - Scheduling Conference

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Rule 37 (f) - Safe Harbor?

Electronically stored information. Absent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide electronically stored information lost as a result of the routine, good faith operation of an electronic information system.



Preservation Summary

- Duty to Preserve
- Complying with Litigation Holds
 - Corporate Counsel Role Widened
 - Document Retention Requires More Attention
- Pull & Park Phenomena
 - Backup Tapes
 - FRCP - Safe Harbor - Rule 37
- Identify Not Reasonably Accessible Data



No more “I won’t ask if you don’t ask”

- **Electronic documents were discoverable under existing case law anyway, but....**
- **The new amendments remove any doubt**
- **... and all excuses.**
- **Now you do not dare “not ask.”**

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



FRCP Effect on Collection

- **Practical**
 - **Cannot save the world of data - relevancy**
 - **Integrate with document management system, data archives, litigation hold process**
- **Closer relationship between Legal/IT**
 - **30(b)6 depositions of IT**
 - **Where is the data**
- **Ordinarily maintained**
- **Reasonably Accessible**

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Not Reasonably Accessible Data

- **Rule 34. Production of Documents, Electronically Stored Information, and Things and Entry Upon Land for Inspection and Other Purposes.**

(a) SCOPE. Any party may serve on any other party a request (1) to produce and permit... to inspect, copy, test or sample, any designated documents or electronically stored information... stored in any medium.

- **Proactive Sampling v. Requesting Party Sampling**

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



FRCP Effect on Privilege

- Define the Process - Rule 26(b)(5)(B)
- State It Up Front
 - Quick Peek - turn over data quickly
 - Claw Back Procedures
- Follow the Procedure
 - Return
 - Sequester
 - Destroy
- Priv Log - how flexible is it?
 - Be Descriptive

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Rule 26. General Provisions Governing Discovery; Duty of Disclosure

(b) Discovery Scope and Limits.

(5) Claims of Privilege or Protection of Trial Preparation Materials.

(B) Information produced. If information is produced in discovery that is subject to a claim of privilege... the party making the claim may notify any receiving party of the claim and its basis. After being notified, a party must promptly return, sequester, or destroy the specified information ... and may not use or disclose it until the claim is resolved.

This is the “claw-back”.



FRCP 34 [amended, continued]

● Rule 34. Production of Documents, Electronically Stored Information.

(b) PROCEDURE. ... the request may specify the form in which the electronically stored information is to be produced....

...[the responding party may object] to the requested form, stating the reasons, and the form it intends to use [instead].

(ii) if a request does not specify, the responding party must produce in the form it is *ordinarily maintained* or in an *electronically searchable form*.



FRCP Changes for Production

- Native
 - Meta-data and embedded data
 - Unique ID
- FRCP Default - Rule 34
 - Requesting Party Specifies or Parties Agree
 - Nothing Specified or agreed to
 - Ordinarily Maintained
 - Reasonably Usable
 - Not More than One Format
- Shared Database

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



FRCP Key Changes

- Introduces Term
 - “Electronically stored information” (ESI)
- Different requirements for paper and electronically stored information
 - Introduces concept of accessibility
 - Shift of burden to responding party for ESI
 - Inadvertent production of privileged material (“claw-back”) R.26(b)(5)(B)
 - Limiting sanctions for spoliation if producing party shows good faith efforts (“safe harbor”) R.37(f)
- Dramatic shift from prior practice
 - Legislates meeting and discussing plan for ESI

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Key Concerns

- Uninformed bench
- Uninformed counsel
- Minimal precedent & guidance
 - Highly technical area with little clarity
 - Intangible subject matter
- Orders of magnitude increases in information
- How evolved is your data analysis and review technology? The other side's?
 - May change depending on the matter

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Processing

- More data than ever
- How to Limit the Dataset
- New Methodologies for Processing
 - Culling Methods
 - Statistical Methods
 - Linguistic Methods

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Sedona Conference Search and Information Retrieval, Principle 1:

In litigation... where the volume of discoverable electronically stored information is large, it may not be feasible to perform human review of every document for responsiveness or privilege, and automated search and information retrieval methods and tools may be necessary and valuable.

*This isn't just a brainstorm
of words and phrases.*

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



“exponentially greater volume”

The discovery of electronically stored information raises markedly different issues from conventional discovery of paper records. Electronically stored information is characterized by exponentially greater volume than hard-copy documents. Commonly cited current examples of

From page 22 of the Commentary by the Rules Committee, Sept 2005

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



How Electronic Collections Are Different

- *Much larger data collections, but that's not all*
 - Meta data - Each item contains much data about itself, both hidden and visible
 - Meta-data is your friend
 - And there are *many* very small items
- Duplicates - Collections are often very largely redundant
- Mixture of informal and formal, personal and professional, highly useful and highly useless content, cryptic content
 - Email has replaced paper
 - Email has replaced phone calls

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



How to Limit the Dataset for Review

- Limit custodians
- Date ranges
- Duplication
 - Vertical,
 - horizontal,
 - Entire set; Each time data is added
 - Know your vendor's methodology

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Culling Methods

- Keyword Culling
 - Stemming
 - Boolean Connectors
 - Synonyms
- Query by Example
 - Find Me More of These Type Documents
- Not for Categorization of Data
 - Use Culling to find something specific

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Categorization Methodologies

- Statistical
 - Clustering
 - Fuzzy Logic (assigning values to things - tall, rich, living in San Diego)
- Linguistic Methods
 - Ontologies

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Why Categorize Dataset?

- Responsive
 - Per issue
 - Foreign language segregation
- Non-responsive
 - Junk - porn, SPAM
 - Private
 - Non-responsive business info
- Privilege
 - Over designate
 - More concentrated review

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Categorization Benefits

- Concentrate Effort
 - What is a Concern?
 - Certain Custodians
- Use Resources Wisely
 - Antitrust expertise
 - Foreign Language
 - Initial Review is Complete with Categorization

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



New Methodologies

- Automated Review
 - Statistical Analysis, Testimony
 - Defensible Process
- Advanced Analytical Techniques
 - Data analysis
 - Graphic Depictions

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Budgets

- Be Thorough
 - All Vendors
 - All Lawyers
- See the Checklist for costs
- Consider 2 Biggest Costs
 - Collection/Processing
 - Review

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Budget Considerations

- Single v. Multiple Vendors
- General Fees
 - Set-up
 - User License Fees
- Training Costs
- Bundling

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Processing to Production Costs

- Data Loading
- Processing Costs
- Hosting Costs
 - Shared Databases
- Archiving Data
 - For later use
 - Data Decisions
- Production Costs

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Other Costs

- Professional Services
- Termination Costs
 - Transfer Out
 - Conversion of Data
- Service Level Agreement

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Kay N. Pfeiffer

- Review (including Privilege)
- Form of Production
- Closing the Matter
- Documentation

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Document Review (Including Privilege)

- Planning
 - Define timeline, approach, staffing, review criteria, use needs, database functionality, production specs
- Staffing
 - In-house and outside counsel, service providers, subject matter experts

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Document Review (Including Privilege)

- Method
 - Human, Computer Assisted, Computerized, Combination
- Format
 - Online Native Software, Native Using Quick Viewer, Online Image, Blowback Hardcopy
- Structure
 - Database, Coding and Comment Fields, Log Related Fields

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Document Review (Including Privilege)

- Criteria
 - Responsive (yes/no)
 - Case Specific Coding
 - Privilege (yes/no)
 - Reason (Drop Down Menu)
 - To populate Privilege log
 - Redaction Needed
 - Complete (yes/no)
 - Confidentiality Designation
 - Undetermined Option

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Risks of a Poorly Planned Review

- Failure to find responsive documents
- Failure to recognize responsive documents
- Failure to recognize privileged documents
- Inconsistent treatment of documents (e.g., duplicates)
- Failure to complete project in a timely manner

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Production Options

- Form of Production
 - Paper
 - Electronic
 - Image – PDF/TIFF
 - With Search Tool
 - Native
- Media Selection
 - CD/DVD/External Drive
- Timing: Plan Ahead – Work Backward to set the Schedule

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Production Options

- Pros and Cons of Forms of Production
 - Native
 - Active, searchable document file (+)
 - Cannot be redacted (-)
 - Dynamic, hard to maintain integrity (bates numbers, edits) (-)
 - Metadata maintained (+)
 - Image
 - Cannot be searched without text file (-)
 - Can be redacted (+)
 - Stable, easy to maintain integrity (+)
 - Metadata not maintained (-)

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Closing the Database

- Legal Hold & Record Retention Compliance
- Opportunities for Re-Use

---versus---

- Risks of Maintaining
- Storage Costs

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Documentation of Investigation

- Substantiates Reasonableness of Effort
- Format
- Consistency

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Documentation of Investigation

● Purpose

- Provide Documentation of Reasonable/Good Faith Efforts
- Protect Company Interests
- Provide Information to Explain the Negative
- Facilitate Case Information Management
- Satisfy Amendments to FRCP
- Avoid Rework
- Provide Historical Investigation Data to Future Case Teams

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Documentation of Investigation

● Design of Tool

- Excel Spreadsheet with Advanced Functionality
- Automated Formatting Based on Company Records
- Ability to Generate Reports Base on Populated Data
- Build in Flexibility

● Benefits of Standard Tool

- Consistency
- Accuracy
- Defensibility

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Documentation of Investigation

- Investigation Documentation may be Discoverable
- Anticipate Production
 - Track Data without Commentary
 - Update Chart as soon as Practical after Activity to Ensure Accurate Data
 - Track All Investigation Activity

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



Preservation Process

The new rules do not touch substantive law on the duty to preserve.

- Typically duty to preserve arises before you know much about the data, or the case.
- Get out notices and get back acknowledgments.
- Ask clients only to preserve, not to collect or review.
- Instruct IT to suspend tape rotation and other SOPs.
- Possibly make litigation-specific backups.

ACC's 5th Annual Corporate Counsel University: New Challenges/New Solutions

April 29-May 1, Hyatt Regency St. Louis at Union Station



What does that Report look like?

Form 35

The Committee recommends conforming changes in Form 35, the parties' report to the court of their discovery plan.

Form 35. Report of Parties' Planning Meeting

- 1 *****
- 2 **3. Discovery Plan.** The parties jointly propose to the court the
- 3 following discovery plan: [Use separate paragraphs or
- 4 subparagraphs as necessary if parties disagree.]



- 38 FEDERAL RULES OF CIVIL PROCEDURE
- 5 Discovery will be needed on the following subjects:
- 6 _____ (brief description of subjects on which
- 7 discovery will be needed)_____
- 8 Disclosure or discovery of electronically stored
- 9 information should be handled as follows: _____ (brief
- 10 description of parties' proposals) _____
- 11 The parties have agreed to an order regarding claims of
- 12 privilege or of protection as trial-preparation material
- 13 asserted after production, as follows: (brief description
- 14 of provisions of proposed order)
- 15 All discovery commenced in time to be completed by
- 16 _____ (date)_____. [Discovery on _____ (issue for
- 17 early discovery) _____ to be completed by
- 18 _____ (date)_____.]

Custodian Box Code Office Location Phone Number	Office Collection Date and Reviewer(s)	Certification Completed and Collected	No. of Documents Collected (Boxes)	Date Boxed	Return Date	Removable Storage	Additional sources of Documents (Requires Follow Up)	Completed	Comments/Investigation Status	Office completion date

Contact History					
<i>Userid</i>	<i>Name</i>	<i>Phone</i>	<i>Title</i>	<i>Date Contacted</i>	<i>Response</i>

Activity Log		
Task	Actions Taken	Status

File Share Data							
Server	Share	Location	Pull Data	Date Request Sent	Date Request Received	Volume (Size in MB)	Comments

Custodian List for Electronic Collection

Custodian Name	Location	Status	Hard Drive	Volume (Size in MB)	.pst	Volume (Size in MB)	e-Mail	Volume (Size in MB)	Date Request Sent	Date Request Received	Comments

Hold Documentation

Name	Hold Letter Number	Hold Letter Name	Date Hold Notice Sent	Comments

Interviews				
Name	Interview Date	Interviewer(s)	Interview Notes (Attach File)	Comments

Testimony					
UserID	Name	Type	Date(s) of Testimony	Matter # and Name	Comments

Sales Record Analysis								
<i>Year</i>	<i>Customer (Sold To)</i>	<i>Customer (Ship To)</i>	<i>Product Code</i>	<i>Product</i>	<i>Invoice #</i>	<i>Invoice Date</i>	<i>Volume</i>	<i>Dollars</i>

Contact List									
<i>Company</i>	<i>Name</i>	<i>Address</i>	<i>City</i>	<i>State</i>	<i>Zip</i>	<i>Phone</i>	<i>E-mail</i>	<i>Cell Phone</i>	<i>Fax</i>