

THE E-DISCOVERY PROCESS

John Murdock, Esq.
Epstein Becker & Green, PC

Scope of this presentation

How to best address client concerns about the E-Discovery process;

The duties and standards for E-Discovery; and

How duties are performed and standards met.

PART I

Client concerns

11/13/2007

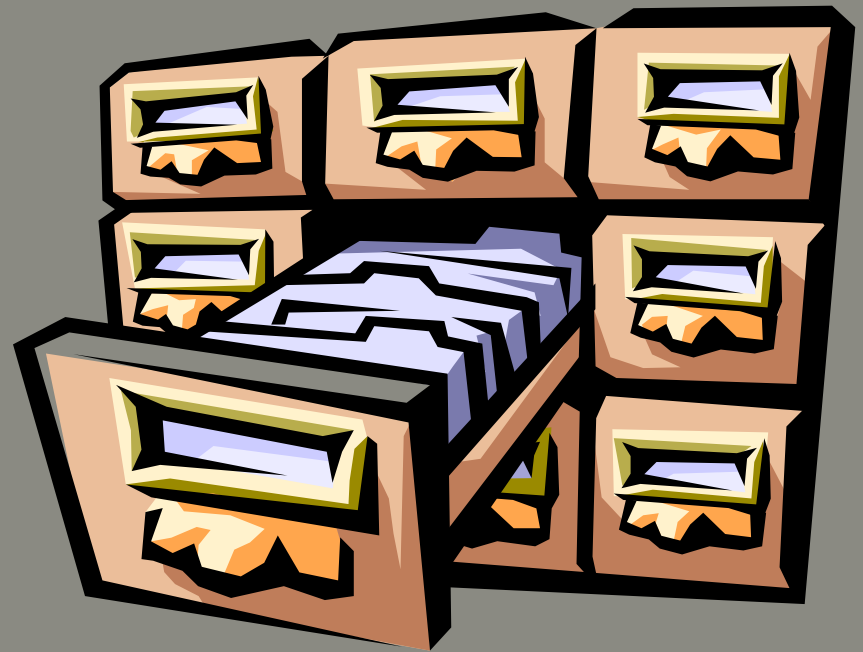
Most information generated today starts in DIGITAL form!

- Almost every document that you receive or produce in discovery, originated in digital form, and paper copies are simply incomplete versions.

11/13/2007

Where is information stored?

- At least 70% of electronic documents are never printed
- If you limit consideration to only paper documents, it is the equivalent of managing only 3 of 10 file drawers of potentially relevant information.



11/13/2007

How much information?

- ① 1GB = 70 feet of paper = 300 Reams of paper.
- ② 20GB = 1400 feet of paper = Empire State Building
- ③ Not uncommon to find a minimum of 80GB in a laptop.
- ④ Just came out with consumer level 750GB and 1 terabyte hard drives!

11/13/2007

The E-Explosion

- Today an estimated 60 billion emails are sent per day...
- In 2006, corporations were expected to generate well over 17.5 trillion electronic documents annually.
- 25 billion messages are exchanged by office workers every day
- Email generates 400,000 terabytes of new information each year
- The average corporate user experiences at least 852 emails a month
- 1 in 20 companies have battled a workplace lawsuit triggered by email

11/13/2007

Typical contexts in which attorneys perform E-Discovery

- Responding to discovery requests in litigation;
- Responding to investigative subpoenas;
- Conducting due diligence for business transactions;
- Internal investigations;
- Complying with agreements, e.g. corporate integrity agreements requiring continuous reporting of information; and
- other

Typical client concerns about the E-Discovery process

Unnecessary	<ul style="list-style-type: none"> • overly broad or poorly defined date and subject matter parameters of information to be produced
Burdensome	<ul style="list-style-type: none"> • IT and other employee time
	<ul style="list-style-type: none"> • electronic information system time and invasion
	<ul style="list-style-type: none"> • employee complaints
Threatening	<ul style="list-style-type: none"> • disclosure of business information or information that may prompt further inquiry by third parties
Expensive	<ul style="list-style-type: none"> • often 50% to 80% of attorney fees for a project
	<ul style="list-style-type: none"> • involvement of experts
	<ul style="list-style-type: none"> • involvement of vendors
Uncontrolled	<ul style="list-style-type: none"> • takes too long
	<ul style="list-style-type: none"> • process is unclear
	<ul style="list-style-type: none"> • unanticipated and invisible activities and costs
	<ul style="list-style-type: none"> • scope and sources of information to be produced expand

12/14/2007

Typical distribution of costs in matters involving E-Discovery

-
- Hearings
 - Meetings and negotiations with opposing counsel
 - White Paper/Briefs/Written opinions
 - Development of strategy
 - Review of key information
 - Legal research and analysis

20% to 50% of costs

- Selection, gathering, review and production of hard copy documents
- Production of electronically stored information (“ESI”)
- Review of ESI
- Processing of ESI for review
- Identification and extraction of ESI
- Coordination of client preservation of ESI
- Assessment of key electronic system functionality
- Identification of key electronic system information systems
- Identification of key electronic system control personnel

50% to 80% of costs

Our approach to effectively addressing client concerns about E-Discovery should be informed by understanding

The purposes of E- Discovery

Our duties and the standards of practice for information discovery complaint with the E-discovery provisions in the Fed. R. Civ. P.

The process of E-Discovery

PART II

Duties and standards

11/13/2007

Purpose of E-Discovery

- “The purpose of discovery is to provide a mechanism for making relevant information available to the litigants. ‘Mutual knowledge of all the relevant facts gathered by both parties is essential to proper litigation.’”
 - Fed. R. Civ. P. 26 Advisory Committee Notes (1983), quoting *Hickman v. Taylor*, 329 U.S. 495, 507 (1947).
- A corollary is that we want to produce:
 - Information that is a relevant authentic record reliable for admission into evidence or acceptable by a third party.

Relevance of Fed. R. Civ. P.

The Federal Rules of Civil Procedure are relevant to E-discovery in a federal court litigation but they should be considered even in a non-litigation context

- ESI production disputes in a non-litigation context that the parties do not resolve themselves, will, inevitably, be referred to courts and tribunals that are accustomed to the standards of practice compliant with the Fed. R. Civ. P. and analogous rules of procedures in State courts for resolving ESI disputes.

Examples of circumstances in which a court may become involved in what was a non-litigation matter involving ESI production

Context	Issues
Investigative Subpoenas	<ul style="list-style-type: none">• motions to quash or limit or compel• form in which information to be produced• Burdensomeness and cost shifting• sampling and testing• onsite inspection of electronic systems• providing support and instruction to party receiving production
Due Diligence, Internal Investigations and Other Productions By Agreement	<ul style="list-style-type: none">• allegations of misrepresentation because of failure to disclose relevant information; allegations of attorney negligence
All of the above	<ul style="list-style-type: none">• privilege and work product waiver• allegations of spoliation

Mindfulness of the Fed. R. Civ. P. E-discovery provisions makes it more likely than not that our ESI production practices will withstand legal attack and judicial scrutiny, produce information deemed complete, authentic, and sufficient, avoid potential problems, e.g. waiver and spoliation, thus serving the clients' interests.

Electronically stored information

Definition of “electronically stored information” (“ESI”) – Fed. R. Civ. P. 34(a) 2006 Amendment Advisory Committee Notes describe “electronically stored information” as “...any type of information that can be stored electronically.”

Core duties and standards created by or that follow from the E-Discovery provisions of the Fed. R. Civ. P

1. Reasonable understanding of the information systems used by the client – See Fed. R. Civ. P. 26(f); 2006 Amendment Advisory Committee Notes to subdivision (f).
 - N.b. Where organizations store and how they organize electronic information is idiosyncratic.
 - Rule 26(f) requires parties to confer and plan for discovery, including preservation of discoverable information.

2. Promptly assess client's electronically stored information – See Fed. R. Civ. P. 26(a)(1); 2006 Amendment Advisory Committee Notes

- Rule 26(a)(1) requires that parties make initial voluntary disclosures of certain categories of information, including a description by category and location of “all” ESI that may be used “to support its claims or defenses”.

3. Advise and assist client in, and perform an adequate assessment of client's preservation of its electronically stored information – See Fed. R. Civ. P. 26(a)(1), and 37(f); 2006 Amendment Advisory Committee Notes to subdivision (f)

- Distinction between “retention” and “preservation”
- Rule 37(f) creates a “safe harbor” from failure to produce electronically stored information “lost as a result of the routine, good-faith operation of an electronic information system.”
 - What is “good faith”?
 - What is “routine operation”?
- Implications of any existing preservation orders, or statutory, or regulatory, or contractual obligations to retain or preserve information subsequently lost on Rule 37(f) analysis.

-
4. Appropriately distinguish between ESI that is reasonably accessible from that which is not – See Fed. R. Civ. P. 26(b)(2); 2006 Amendment Advisory Committee Notes to subdivision (b)(2)
- Rule 26(b)(2) creates “limitations” on discovery which include permitting a producing party to refuse to produce electronic information that is “not reasonably accessible” absent requesting party making a showing of “good cause” to compel production.
 - “Not reasonably accessible” information must be identified in a privilege log provided to opposing party.
 - N.b. Do not access, even to sample or test, ESI thought to be “not reasonably accessible” and agreement with the opposing party or a direction from the court because such access may result in a waiver of the “not reasonably accessible” limit waived.

5. Perform a privilege review of electronic information to be produced that is reasonable in the circumstances – See Fed. R. Civ. P. 26(b)(5)(B); 2006 Amendment Advisory Committee Notes to subdivision (b)(5)(B)

- Rule 26(b)(5)(B) permits a party that has produced privileged information in discovery to demand its return and requires the receiving party to return the information, but allows the receiving party to petition the court to determine whether the information is privileged and whether the privilege has been waived.
- By contrast, a “clawback” agreement may describe what information the parties deem privileged and which, if disclosed in a production, must be returned without challenge and as to which between those parties there will not have been a waiver of privilege.
- A “clawback” agreement is likely of no effect against a third party’s assertion of privilege waiver from the disclosure. To have any chance at effectiveness against third party, a “clawback” agreement should be embodied in a court order.

6. Consider and discuss with opposing counsel as appropriate

- Scope of preservation obligation -- See Fed. R. Civ. P. 26(f); 2006 Amendments Advisory Committee Notes subdivision (f)
- Form in which electronically stored information is to be produced – See Fed. R. Civ. P. 26 (f) and 34(b); corresponding 2006 Amendments Advisory Committee Notes subdivision
- Types of meta-data that must be produced
- Manner in which information is to be marked for production
- Manner in which information is to be redacted
- Procedures for post-production assertion of privilege or protection of information
- Time period for which discovery is to be sought
- Sources of a party's searchable electronically stored information that is reasonably accessible
- Search methodologies to be used, including search terms or filters
- Burden and cost of retrieving and reviewing electronically stored information that is reasonably accessible
- Interviews with information systems personnel
- Need for protective orders or confidentiality orders
- Value of staged discovery

Document performance of the duties

- Documenting steps taken to identify, preserve, retrieve, and produce clients' ESI
- Litigation hold letter
- ESI and electronic systems letter and questionnaire
- Documentation from client of preservation protocol
- Notes of interviews with relevant IT personnel and ESI custodians
- Documentation of client and/or ESI vendor searches and extraction of client ESI from client, e.g. chain of custody documents

Information Hold Letter

Identifies the persons who are likely to have relevant information (preliminary) and communicates a preservation notice to those persons;

Communicates the preservation notice in a manner that ensures the recipients will receive actual, comprehensible and effective notice of the requirement to preserve information;

Is in written form;

Clearly defines what information is to be preserved and how the preservation is to be undertaken;

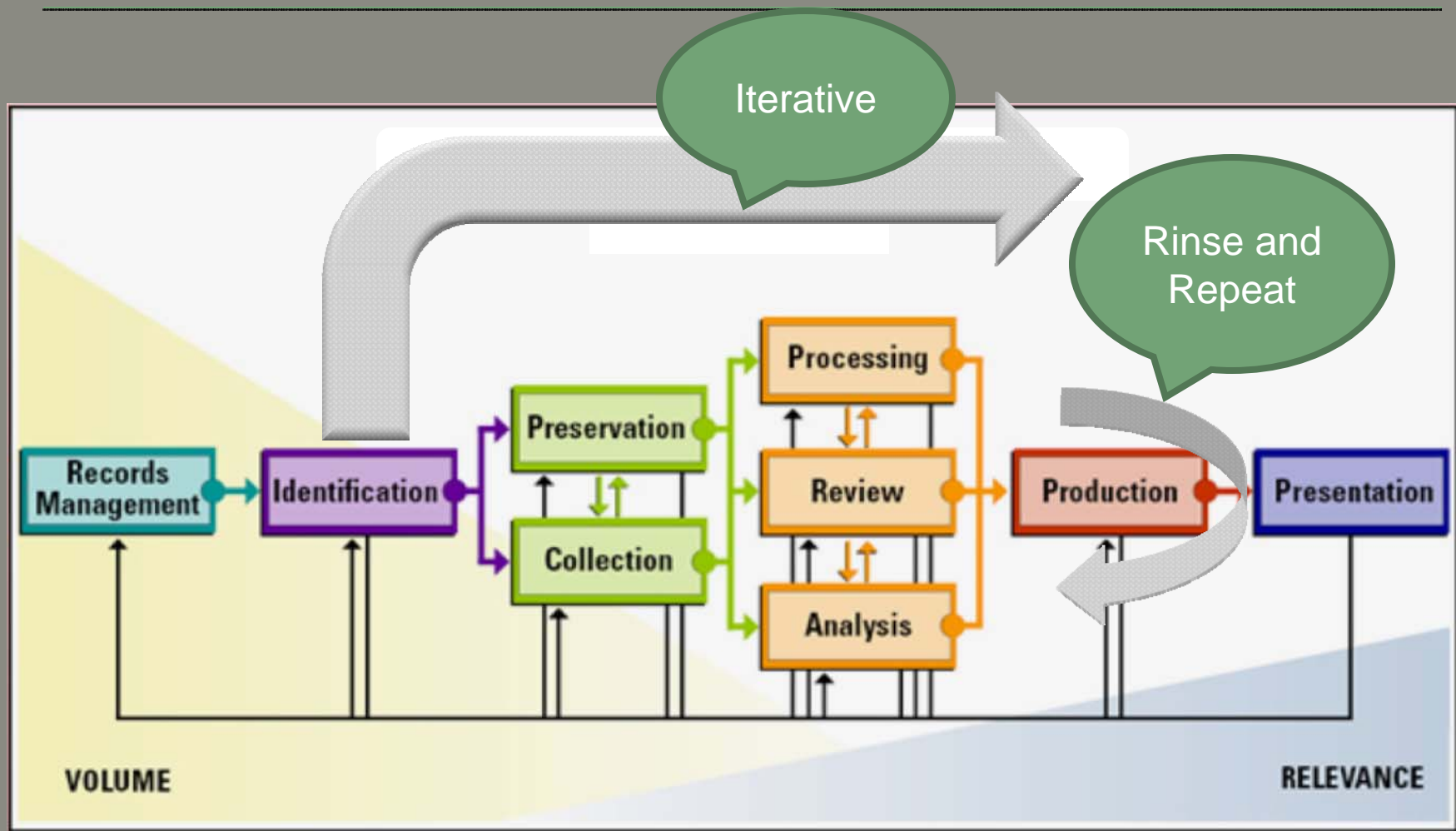
Is periodically reviewed and, when necessary, reissued in either its original or an amended form.

PART III

Process

12/14/2007

Discovery Process Overview



12/14/2007

Identification

The process of learning the location of all data which you or your client may have a duty to preserve and potentially disclose in a (pending or prospective) legal proceeding.

Common Data Sources

Local Area Server for
home office

Personal Share or
Personal Folders on
Server

Dedicated Server for
[XYZ, Inc.]

Laptop and/or Office
Computer

Home Computer,
Blackberry and/or
PDA

E-mail, including
archived e-mail and
sent e-mail

E-mail Trash Bin,
Desktop Recycle Bin

Removable Storage
Media, such as
disks, CDs, DVDs,
memory sticks, and
thumb drives

Office Files

Personal Desk Files

Files of any
Administrative
Personnel

Files located at
home

Digital evidence resides here



11/13/2007

And here ...



11/13/2007

And here!



11/13/2007

Digital information comes in many forms

Examples:

- Typed documents and correspondence
- Email exchanges
- Spreadsheets
- Data bases
 - Financial / Corporate / Medical / Personal
- Previously viewed web pages
- Metadata
- Short Message Service (SMS)
- Appointment calendars and “To Do” lists
- Lists of stored telephone numbers

11/13/2007

Metadata – Data About Data

- File Dates:
 - Created
 - Modified
 - Accessed
 - Printed
- Title
- Subject
- Author
- Manager
- Company
- Last Saved By (name)
- Revision Number
- Total Editing Time
- And over 25 other document attributes

11/13/2007

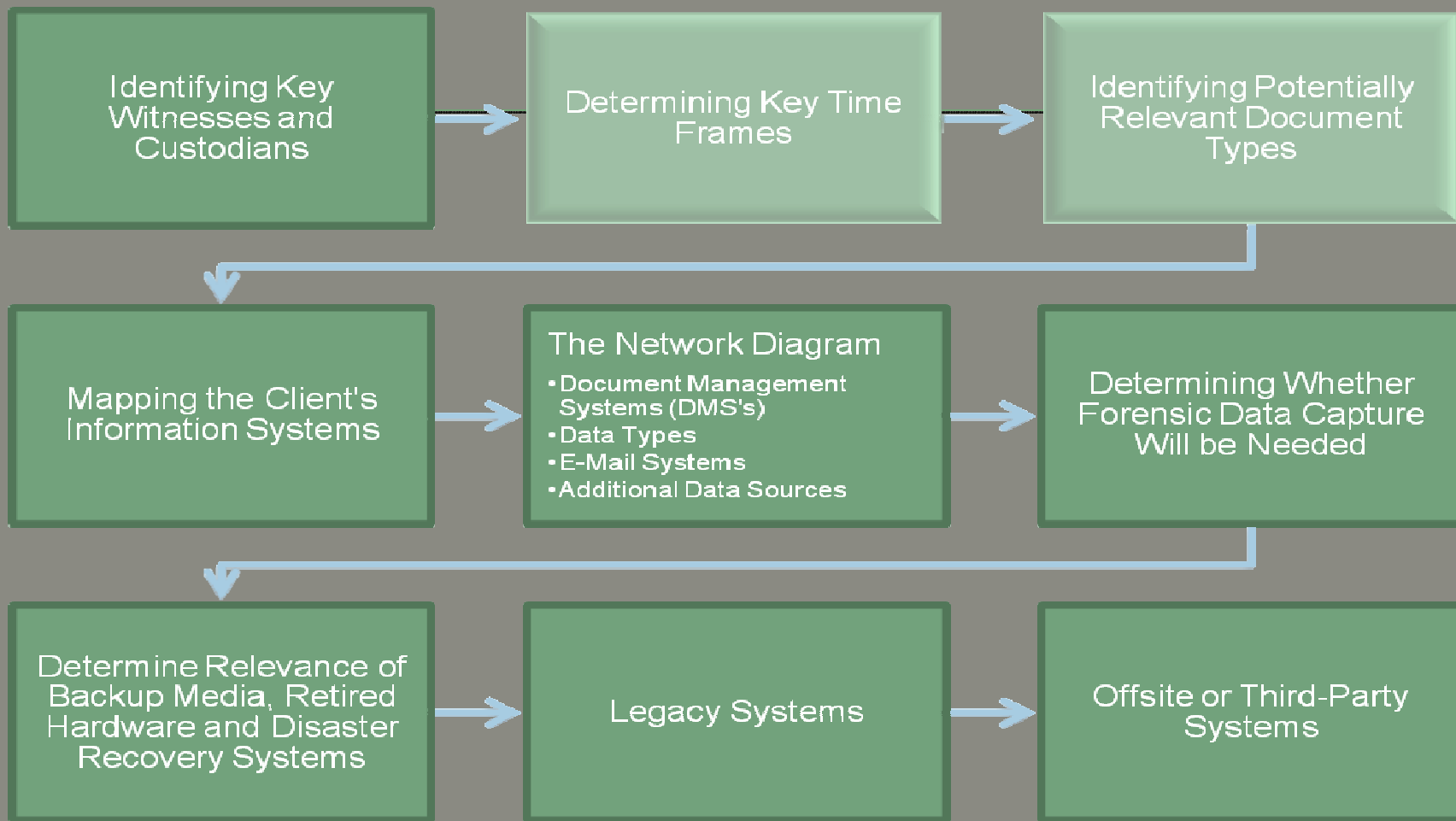
Identification Process

Potential custodians should be interviewed.

Begin with key players and business information systems.

Inventory client information systems, map relevant information systems to key custodians and business information.

Consider: email, backup tapes, group shares, knowledge management systems, home computers, personal digital assistants.



Data Silos

Online/Production Data

- Data in currently running production systems, including e-mail, databases, commercial off-the-shelf (COTS) applications, or other active company records.

Offline Data

- Files stored on network file shares, local desktop or laptop file systems, on portable storage such as CDs or DVDs, on portable storage devices, on external hard drives, in Personal Storage Files such as a PST file.

Archive Data

- Files stored in a corporate records management system or within an archive including e-mail and instant messaging data.

Backup Data

- Files stored on backup media of any sort, including tapes, snapshots, file-based backups, backups of portable storage devices in any location (onsite, offsite, in transit, at employees' homes, or awaiting disposal/re-use).

Understanding Digital Forensics

- Digital devices are used in all types of businesses for many different purposes. Digital Forensics is often employed to determine the existence, or absence, of information thought to be relevant to an inquiry.

Collection

The acquisition of electronic information (data) identified as potentially relevant in the identification phase.

Questions Always Precede Action

- Just a few examples of serious questions that need to be asked (and answered) before starting.

- What are you looking for?
- Why are you looking for it?
- Who owns the digital device?
- Where is the digital device?
- Who uses the digital device?
- Is the digital device on a network?
- Do you need a court order?
- Does the person who uses the digital device have a reasonable expectation of privacy?
- If owned by a company or other entity, is there a published policy in place that informs users that the digital devices they use may be searched without warning, and they do not have any expectation of privacy while using it?
- How do we gain access to the digital device, and when?

Chain of Custody Records

Regardless of the collection method employed, strict chain of custody records must be maintained.

A unique media ID - This becomes the core tracking number and all of the information extraction from the media;

Date and time of receipt or collection of the evidence;

The name of the person(s) collecting and/or taking possession of the evidence;

A description of the type of evidence (8mm tape, hard drive, etc.);

A description of what the evidence represents (Exchange server-e.g. Chicago);

Any label information (exact);

Serial numbers;

Description of the physical location and custodian at the time of possession;

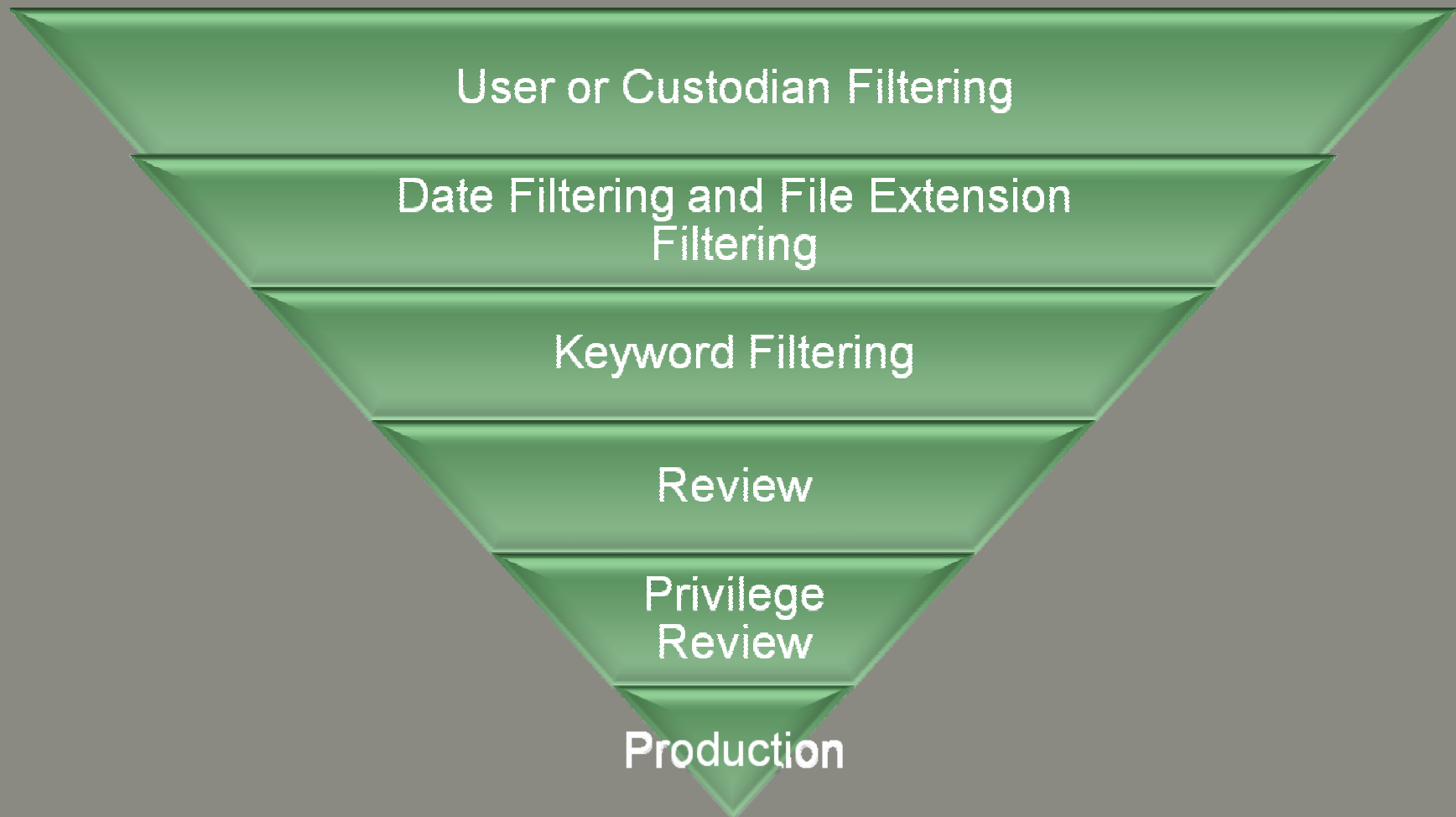
Areas for transferring possession of the media within the collecting organization or to a vendor;

Description of collection methodology;

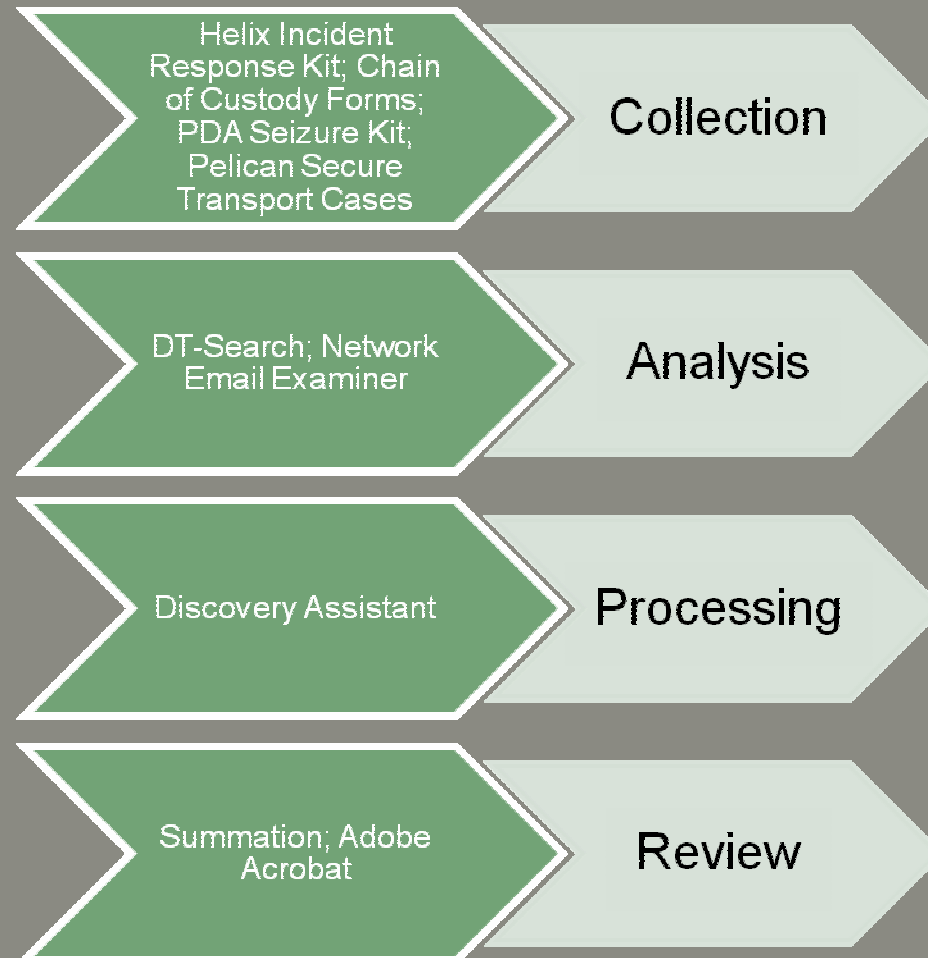
Detailed description of data harvested on site; and

Check lists related to any on-site filtering of data during collection.

Processing, Reviewing, and Analysis




Available Tools



12/14/2007

File Format Characteristics

Formats	Searchable	Metadata	Redaction	Page Numbering	Processing Considerations
Native Files	Yes	Yes	Yes	No	Native Files Usually Associated With Database
TIFF	No	No	Yes	Yes	Usually Extracted Before Tiffing
PDF	Both	Yes	Yes	Yes	Image PDF Can Be OCR'd
PAPER	No	No	Yes	Yes	Usually Scanned Converted to TIFF
Database/ Spreadsheet	Yes	Yes	No	No	Formatting Issues Sometimes Require Native Production and/or Preprocessing (e.g. NSF, PST, MDB, SQL)
ASCII/ UnicodeText	Yes	Yes	No	No	A Type of Native File



Search Techniques

- Keyword Search
- Boolean Operators
- Proximity Operators
- Concept Search

Culling and Searching Considerations

- Sampling and Developing a Strategy
- Early Review Benefits Culling and Searching
- Identify the "Potentially Relevant"
- Suppress, Don't Delete
- Measure Success
- Is It Defensible?

Impact of Keyword Selectivity

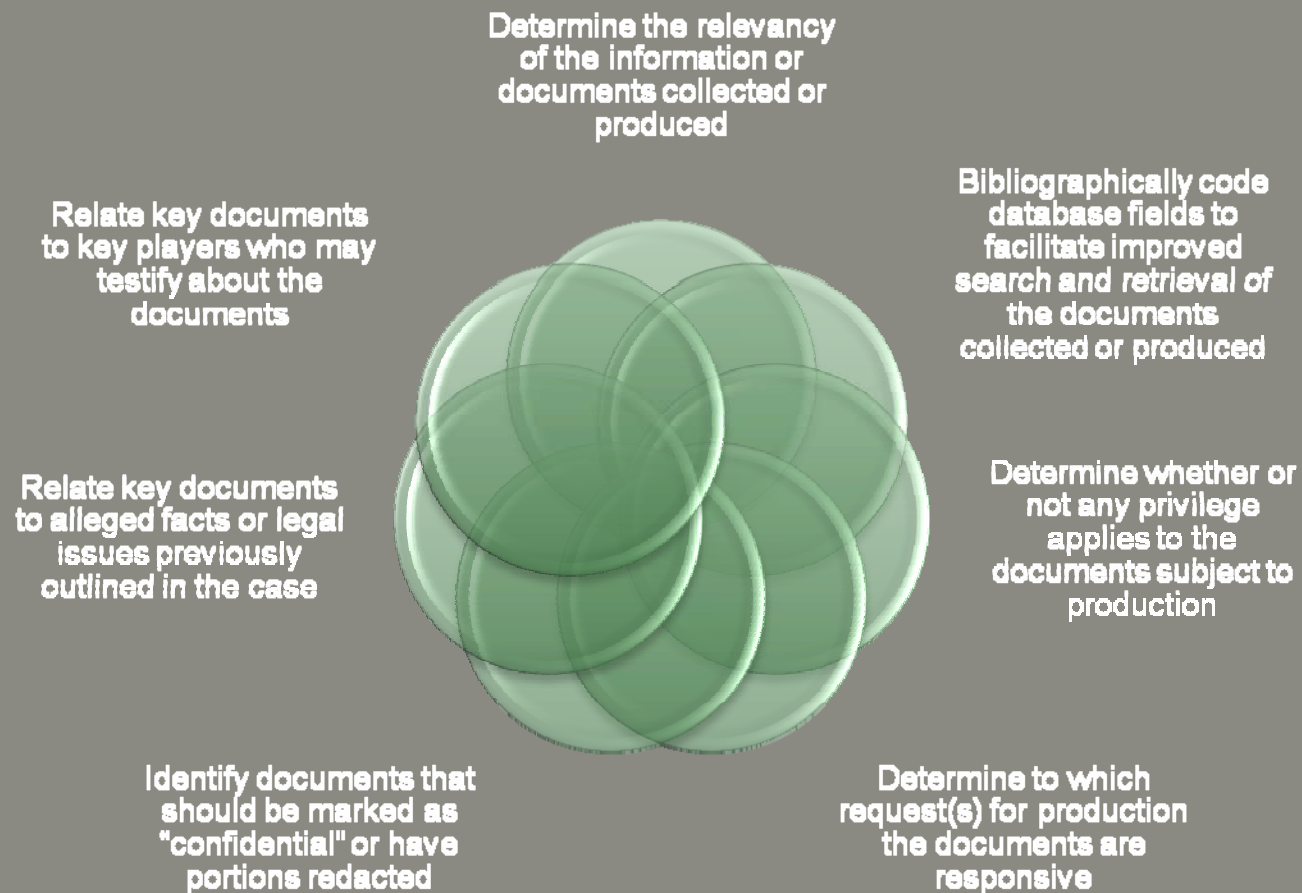
Bernchmark	Relative effect of keyword selectivity		
	Broad Keywords	Median	Highly Selective Terms
Images ^[1] per GB	78,671	47,213	18,534
Images per file email	11	4	2
Images per file app files	63	10	3
Files per GB email	36,530	22,572	9,934
Files per GB app files	20,305	15,791	7,553
GB per custodian	5	2	1
GB per custodian app files	4	1	0

12/14/2007

Quality given keyword selectivity

Culling Rate Percentages			
<i>Deduplication</i>	51%	21%	6%
<i>Searching/Filtering</i>	64%	61%	23%
<i>Non-printable files</i>	22%	5%	2%
Processing Speeds (in hours)			
Process time per GB native	117	33	11
Process time per GB image	35	32	23
Process time to first deliverable	53	35	21
Process time by file type (minutes)	4	3	2
Process time by file type (minutes)	6	4	3
Process time by file type (minutes)	2	3	2
Quality			
First pass quality yield %^[2]	57%	78%	73%

Review Objectives



Defining Review Protocol

Rules for adding additional bibliographic information to the database such as: type of document (memorandum, letter, email), entities named in the document, marginalia found on paper docs that have been scanned, etc.

Publication of look-up table to aid in identification of key players and known entities

Rules for identification and subjective coding/tagging of the responsive and producible documents with an outline of the issues to be identified and possibly sample documents that are representative of the issues at hand

Rules for the handling of documents requiring redaction

Rules for identifying and coding/tagging privilege documents

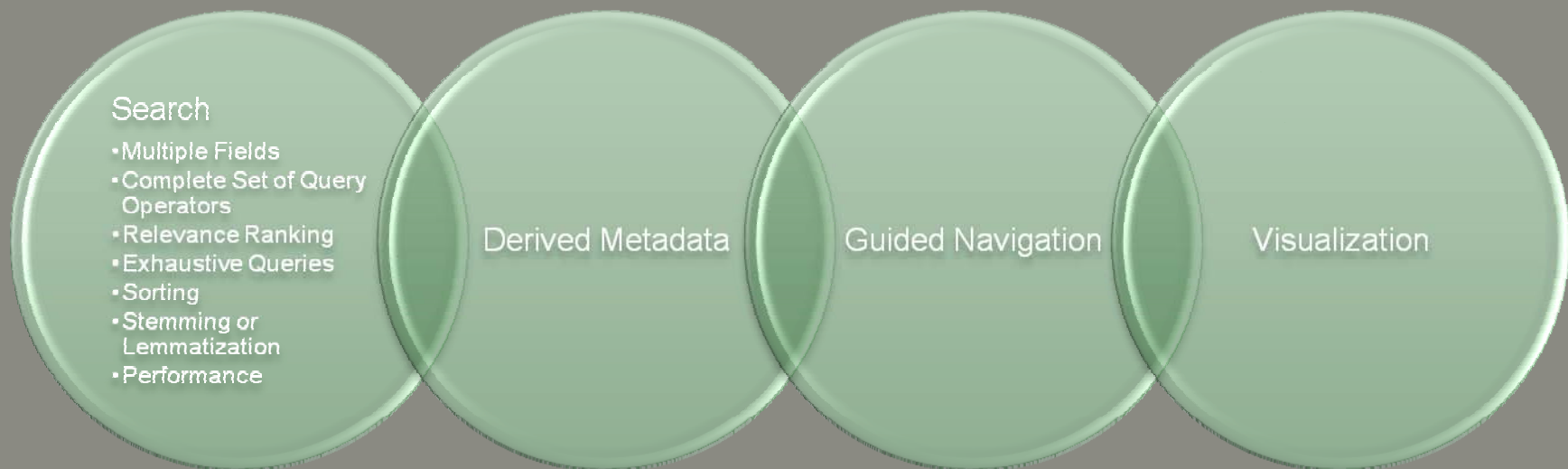
Rules for adding annotations to documents in the form of attorney notes and work product

Rules for coding/tagging emails and their attachments and email threads

Rules for the identification and tagging of non-responsive documents, spam and junk email.

Rules for handling of unreadable, password protected or other faulty documents

Analysis Techniques



Production Considerations

Factors to Consider

- Specific Questions To Consider During Negotiations
- How Will Paper Documents Be Produced?
- What Types Of Electronic Documents Make Up The Data Set?
- What Formats For The Production Documents Provide Access To The Data Necessary To Best Address Issues In The Case?
- What Types of Media Should Be Used To Produce And Receive Production Documents?

Other Considerations

- Production Capabilities and Limitations
- What Technical Formats For The Data Will Be Needed By Each Party?

Metadata

- Using a Forensics Expert
- Meet and Confer
- Considerations for Producing Metadata

Rolling Production

Types of Production

Paper

Quasi-Paper

Quasi-Native*

Native

- Special Considerations for Native Productions
- Where Native Production May be Necessary
- Alteration of Files
- Metadata

Production Formats

Receiving Options	Types	Cost to Organize	Cost to Maintain	Ease of Use	Metadata
Paper	Regular, attachments	High	High	Low	Coded by Collector
TIFF	Type IV	Moderate	Moderate	Low	Extracted During TIFF PROCESS
Native File	Word, Outlook, Visio	Moderate	Moderate	Moderate	Yes
Database/ Spreadsheet	Access, Excel	Moderate	Moderate	High	Yes
Native File Database	Oracle, MySQL	Moderate	Moderate	High	Yes
Litigation Support Database	Summation, Concordance	Moderate	Moderate	Moderate	Yes
Hosted Repository	Lextranet, I-Connect, CaseValut	High	High	High	Yes

*** END ***

11/13/2007