Electronic Discovery

e-Discovery: A Primer

Mauricio Perry, CRM, CEDS
Disclaimer

• I am not a lawyer
• The ideas exposed here are not to be construed as legal advice but are educational information
• Consult your lawyer
Objectives

• Learn the lingo
• Understand how eD works and the risks associated with it
• Understand how a Records Manager can affect it
E-Discovery

IT

Records & Information Management

The Law
“What format shalt I save these in?”
Records

- Record: "information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business". [ISO 15489-1:2001] This definition is independent of the medium of the record (paper, electronic, other objects)
Records Management

• RM is a form of risk management
Disposition of Records

• Why dispose of records?
  – When they are no longer used nor required

• Keeping records for longer than the RRS, exposes the owner to unnecessary risk

• Disposition (usually means destruction) must be done in a way that the record cannot be read or reconstructed: delete all copies of files in every directory/cloud at the same time, shred, etc.
  – Hitting “Delete” is not enough!
Metadata

• Metadata is data about the data. Example: given a Word document, its Properties (metadata) indicates who is the author, when was it created, when was it revised, subject, key words, etc.

• Metadata is very important to for finding electronic documents. Some systems search only in the metadata fields, not in the actual text of the documents. This is the case with scans or photos which are not searchable
Electronically Stored Information

• Where does ESI reside?
  – Desktop and laptop computers
  – Servers
  – External hard drives
  – Flash drives
  – CDs, DVDs
  – Cell phones, mp3 players
  – Tablets
  – Digital Cameras (photo, video)
  – Photocopiers, scanners
  – The “Cloud”
Electronically Stored Information Classification

• ESI sources:
  – Custodial: ESI in possession of one person (current or former employee) (e.g. Word, Excel files)
  – Non-custodial
    • Structured: in databases and wikis
    • Non-structured: in shared or common drives (e.g. Word, Excel files)
Definitions

(Wikipedia)

• The Federal Rules of Civil Procedure (FRCP) govern civil procedure (i.e. for civil lawsuits) in United States federal courts

• Discovery: the pre-trial phase in a lawsuit in which each party [...] can obtain evidence from the opposing party by [...] requests for production of documents, [...] and depositions
Definitions
(Wikipedia and others)

• Legal Hold: preservation of all forms of relevant information when litigation is reasonably anticipated

• Spoliation: destruction or alteration of evidence, usually documents
  – Destruction of records that should be under legal hold constitutes spoliation
  – Spoliation can have grave consequences in a lawsuit

Mauricio Perry, CEDS, CRM
Privilege

- Legal professional privilege protects all communications between a professional legal adviser (attorney) and his or her clients from disclosure without the permission of the client
- Privilege is a complex topic
A New World

Lawyers **must** understand the brave new world of e-discovery (electronic discovery)
Lawyer expelled for five years for telling client to despoil social media content

Disciplinary System Actions

July 2013 – Present

August 2, 2013

Matthew B. Murray, 1852 Wayside Place, Charlottesville, Virginia 22903
VSB Docket Nos. 11-070-088405, 11-070-088422

On July 17, 2013, the Virginia State Bar Disciplinary Board suspended Matthew B. Murray’s license to practice law for five years for violating professional rules that govern candor toward the tribunal, fairness to opposing party and counsel, and misconduct. This was an agreed disposition of misconduct charges.
E-Discovery Reference Model
E-Discovery Reference Model

“Needle in haystack”

Mauricio Perry, CEDS, CRM
E-Discovery Reference Model

Information Management also includes RM + Litigation Readiness
E-Discovery Reference Model

Legal Hold:
1. What to preserve (identification) is based on the data map
2. How/where to preserve
Collection includes Interviews of custodians and Cooperation between opposing counsels, maybe forensic collection too
E-Discovery Reference Model

Successive steps to separate the wheat from the chaff
1. Culling, key words, TAR (e.g. Predictive Coding)
2. Review for relevance and privilege
“Needle in haystack: the bigger the haystack, the more difficult to find the needle”
E-Discovery

• E-discovery can be very expensive if the records are not properly organized, particularly e-mail. This is related to Litigation Readiness

• International e-discovery issues
  – Different countries have different rules
  – Blocking statutes
Zubulake v. UBS Warburg
(Judge Shira Scheindlin)

• Legal counsel must be familiar with e-discovery and the Zubulake decisions
  – The scope of a party's duty to preserve digital evidence during the course of litigation or even when first acknowledged that a chance of litigation exists;
  – Lawyer's duty to monitor their clients' compliance with electronic data preservation and production (litigation hold);
  – Data sampling, so that knowledge about costs and effectiveness of the recovering process are known in advance;
  – The ability for the disclosing party to shift the costs to the requesting party of recovering inaccessible media (backup tapes, for example);
  – The imposition of sanctions for the spoliation of digital evidence.
  
(Wikipedia)
Litigation Readiness

• Simply being prepared for lawsuits at any time
• Keep only necessary ESI, destroy ESI according to procedure and RRS
• Have a constantly updated IT data map: “know what you have and where you have it”. Data maps are one of the first things exchanged during the “meet and confer” between opposing counsels
• List of security and access controls, backup processes, automatic purges, audit trail capabilities, etc.
• Prepare a generic litigation hold protocol
The whole e-discovery process has several stages:

- Collection of ESI
- Processing of ESI (deduplication, de-NISTing, date analysis, format conversion)
- Review for relevancy and privilege (keywords, dates, Technology Assisted Review-TAR/Predictive Coding)
- Hosting
- Production (in agreed upon format, including metadata)
Collection

• It is a joint effort between:
  – The attorney(s)
  – The client contact
  – The client’s IT personnel
  – Forensic collection personnel (if appropriate)
Processing

- The purpose is to eliminate (as much as possible) what is irrelevant to the case but include everything that is relevant

- Culling process
  - Deduplication (when, in what order)
  - Near deduplication
  - De-NISTing (NIST: National Institute of Standards and Technology)
  - Domain filters
  - Date filters

- Once culling process is done, thorough quality control must be performed and documented
Processing
(Kroll Ontrack)
Processing
A Special Case - PDF files

• Composed of multiple files combined or not (Word, Excel, TIFF, JPG, etc.)
• Same elements can be rearranged in different order
• Can have multiple layers, including a metadata layer
• Can be born digitally or scanned

What to do?
• OCR or non-OCR? OCR can alter metadata
• Segregate them and treat them separately
• Use intelligent OCR and/or new OCR technologies
• Convert to PDF/A?
Review

• The purpose is to review the remaining documents for responsiveness, for “key” or “hot”, and for privilege

• Done by lawyers or paralegals
  – Linear review using keywords, date filters
  – Technology Assisted Review (a.k.a. predictive coding computer assisted review)

• Subject to Quality Control done by statistical methods
E-Discovery

Records & Information Management

Statistical Quality Control

The Law

IT
TAR - CAR - Predictive Coding

- Computer technique that uses software that can learn to differentiate what is responsive from what is not. This is achieved by feeding ("seeding") the software documents correctly tagged by experts. Afterwards the software can tag documents accurately.
- Black box
- QC of the resulting sets of documents (e.g. responsive v. non-responsive) by statistical methods (e.g. binomial methods) is a must to be accepted by the courts. Results must be defensible.
TAR - CAR - Predictive Coding

• Note to records managers: this technology can also be used to auto-classify records according to a pre-determined taxonomy
Production

- Files must be produced in **agreed upon** format (paper, native, PDF, TIFF, JPEG, etc., with or without metadata)
- Files may be produced in different media (CD, DVD, hard drive, etc.)
- Files may be hosted in a production database, usually the same for both litigants
Production
Production
(Kroll Ontrack)
Production
(Kroll Ontrack)
A Practical Consideration

Test your search terms **before** reaching agreement with opposing counsel: you do not want nasty surprises!!
Dos & Don’ts
(Kroll Ontrack)

Do:
• Have a document retention policy
• Have an e-discovery plan

Don’t:
• Avoid “hard to deal with” sources of evidence
• Overlook the weight of potential sanctions
• Ignore fast-changing legal landscape
A Court Order Example

• “Both parties have or will immediately arrange to use computer-assisted search technology that permits efficient gathering of documents, de-duplication, maintaining the relationship between emails and attachments, full text Boolean searches of all documents in one pass, segregation or tagging of the search results, and export of all responsive files without cost to the other party.”[...]

• “[a]ll documents in the search result sets shall be produced immediately to the other side in native format including all metadata.”

Cost Factors in e-Discovery

- Volume per custodian
- Number of custodians
- Volume of collection
- Complexity of the case (i.e. of the review)
- Use of 3rd party vendors
Cost Factors: Data Proliferation

Cost of e-Discovery: $1.5 Mil to $3.0 Mil per terabyte

(Osterman Research Study)
Early Case Assessment

• A risk assessment (risk v. benefit) technique to estimate how much will it cost to defend a legal case
Cost Factors in e-Discovery

• Ignorance or lack of budget are not a valid excuse for the courts
e-Discovery in Arbitration

• Some arbitration bodies:
  – American Arbitration Association
  – International Institute for Conflict Prevention & Resolution
  – JAMS
  – NAM

• Each of these arbitration bodies has its own set of rules for the discovery process.
e-Discovery in Arbitration

• It is important to make sure the arbitrators understand e-discovery issues, including its costs
e-Discovery in Arbitration

• Even when arbitration is complete, the duty to preserve may linger on, as documents might be relevant for future litigation
The “Cloud”

- Storage as a service, applications as a service. Examples: Google Drive, Google Docs
- The “cloud” for records and for back-ups. Examples: Yahoo, Amazon, Dropbox
- The danger: loss of control
- Encryption? Where? In original disk? In transit? In the cloud?
- Back-ups in the “cloud”: Where exactly are they?
- Server farms, jurisdictions and destructions
- What happens if the “cloud” company goes bankrupt?
- How fast will they deliver your data during discovery?
- Is special insurance needed?
The “Word Cloud”
Records and Social Media

- Cell phones
- Gmail, Yahoo Mail
- Dropbox
- Twitter
- Facebook
- Pinterest
- Evernote
- Linkedin

And many others: all are potentially discoverable
Discovery of Social Media

• Don’t even think of doing it without expert help
• A screenshot is worth very little (it lacks metadata)
• All sites, all postings are potentially discoverable
Current Trends: Artificial Intelligence

- Predictive Analytics/ E-mail Analytics to detect and prevent future lawsuits

(http://e-discoveryteam.com/2014/03/09/growing-importance-of-non-litigation-services-in-electronic-discovery-law/)
A Final Note

• NSA “touching” e-mails and attachments sent via e-mail may alter metadata
• Its effects on e-discovery is being debated
• Privacy
"Mr. Osborne, may I be excused? My brain is full."
Records Management Related Professional and Industry Organizations

• ACEDS (Association of Certified E-Discovery Specialists): www.aceds.org
• ACESIN (American College of e-Neutrals): www.acesin.com
• AIIM (Association of Information and Image Management): www.aiim.org
• ARMA (Association of Records Managers and Administrators): www arma org
• FRMA (Florida Records Management Association): www.frma.org
• ICRM (Institute of Certified Records Managers): www.icrm.org
• ILTA (International Legal Technology Association): www.iltanet.org
• NAID (National Association for Information Destruction): www.naidonline.org
• PRISM (Professional Records and Information Services Management): www.prismintl.org