

# Electronic Discovery

## ESI 101 (the basics?)

Electronically Stored Information

*010001010101001101001001  
001100010011000000110001*

What to ask for, What to offer, and Why

1

- Dennis Martin
- Litigation Support Analyst – EXO/ISS
- 23 years in Litigation Support, Data Processing, Project Management, DB Administration, and Computer Forensics

ESI 101 (the basics?)

Electronically Stored Information

What to ask for, What to offer, and Why

---

---

---

# Electronic Discovery



2

This CLE is designed around civil litigation, however, the majority of the information presented also applies to criminal cases

Electronic data is relevant in nearly all cases – Bankruptcy, Drug Trafficking, Accounting firm improprieties, Environmental issues, Elderly, Domestic, or Child abuse, etc. can result in a few documents or millions of records to review

And, matters big and small can hinge on a single electronic file, but we need to be able to locate it first

---

---

---

# Electronic Discovery

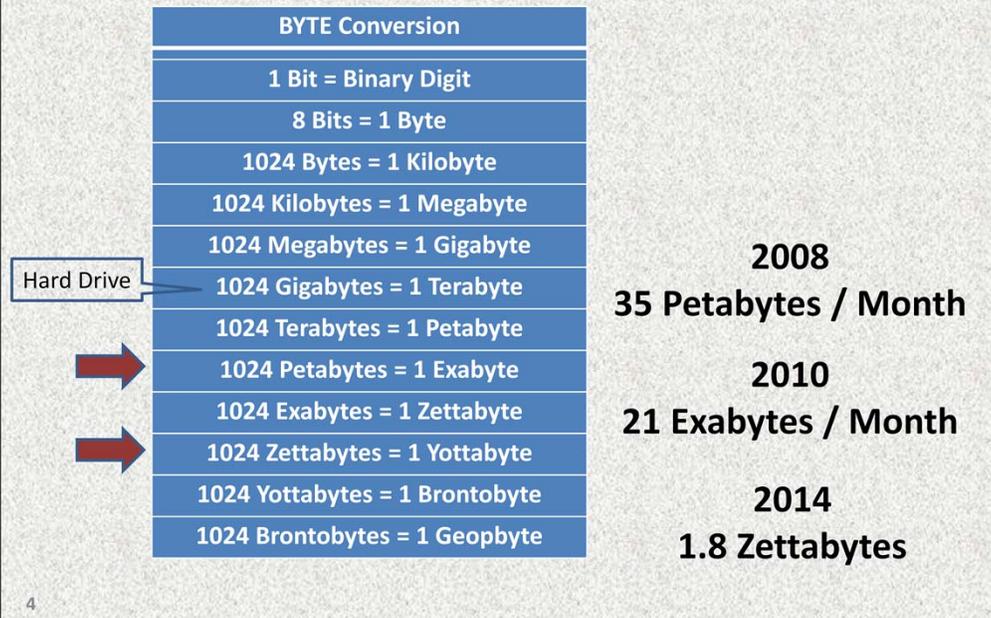
- Email
- Calendars
- Voice Messages
- Instant Messages
- Text Messages
- Word Documents
- Spreadsheets
- Databases
- Internet Logs
- Scripts & Macros
- Deleted Files
- File Fragments
- Image Files
- Audio Files
- Videos
- Diagrams (timelines)
- Config Files
- Executables
- Archives
- Many Others ...

3

## What is ESI:

- **(ESI)** Electronically Stored Information (refer to list)
  - It would be difficult to function without electronic data  
Even many of our homes are controlled by computer systems that may contain stored data
  - Nearly all information generated today is electronic and it is potentially **Discoverable**
  - And, ESI is Growing Exponentially
- 
- 
-

# Electronic Discovery



- A report by The Geneva Association estimated the Global Data Traffic in 2008 at 35 Petabytes per month (35,000 Terabytes)
- A March 2010 study by Cisco Systems estimated **internet traffic alone** to be 21 Exabytes per month (21 million Terabytes)
- A recent study by International Data Corporation (IDC) estimates that early in 2014 there were 1.8 Zettabytes of data on the planet

# Electronic Discovery



5

- Why is this trend so important in litigation?
  - You no longer have a dozen Bankers Boxes to deal with – Now, electronic media may contain the equivalent of hundreds, or even thousands, of boxes of potentially relevant material  
e.g. 12 boxes @ 2500 pages = 1.35 GB (2 CDs) (30K images)
  - We need to be able to sift through these extremely high volumes of, often times, disparate or 'junk' data, requiring modern technologies to assist us
- 
- 
-

## Electronic Discovery

- ABA Rule 1.1  
***“A Lawyer shall provide competent representation to a client...”***
- Maintaining Competence, Comment [6]  
***“To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology...”***

6

- ABA Rule 1.1  
***“A Lawyer shall provide competent representation to a client...”***
  - Maintaining Competence, Comment [6]  
***“To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology...”***
- 
- 
-

# Electronic Discovery



## Sanctions

Awards  
Dismissals  
Adverse Inference

- The ABA Commission on Ethics also stated that the change to comment [6] “**does not impose any new obligations on lawyers**”
- It is only a “reminder to lawyers that they should remain aware of technology...as part of a lawyer’s **general ethical duty** to remain competent”

7

- The ABA Commission on Ethics also stated that the change to comment [6] “does not impose any new obligations on lawyers”
  - It is only a “reminder to lawyers that they should remain aware of technology...as part of a lawyer’s **general ethical duty** to remain competent”
  - Duke Law Journal Vol. 60 → Sanctions for E-Discovery increasing 2005-2009 nearly tripled the number of sanctions
- 
- 
-

## ESI Challenges

- Computers
- Copiers / Printers
- Network Servers
- PDA's
- Phones
- Tablets
- Audio Recorders
- External Drives
- Thumb/Flash Drives
- CD's / DVD's
- Blu-Ray Disks
- Magnetic Discs
- Backup Tapes
- Game Consoles
- TV's
- GPS Devices
- Optical Disks
- Cameras
- Internet / Websites
- Other ...

8

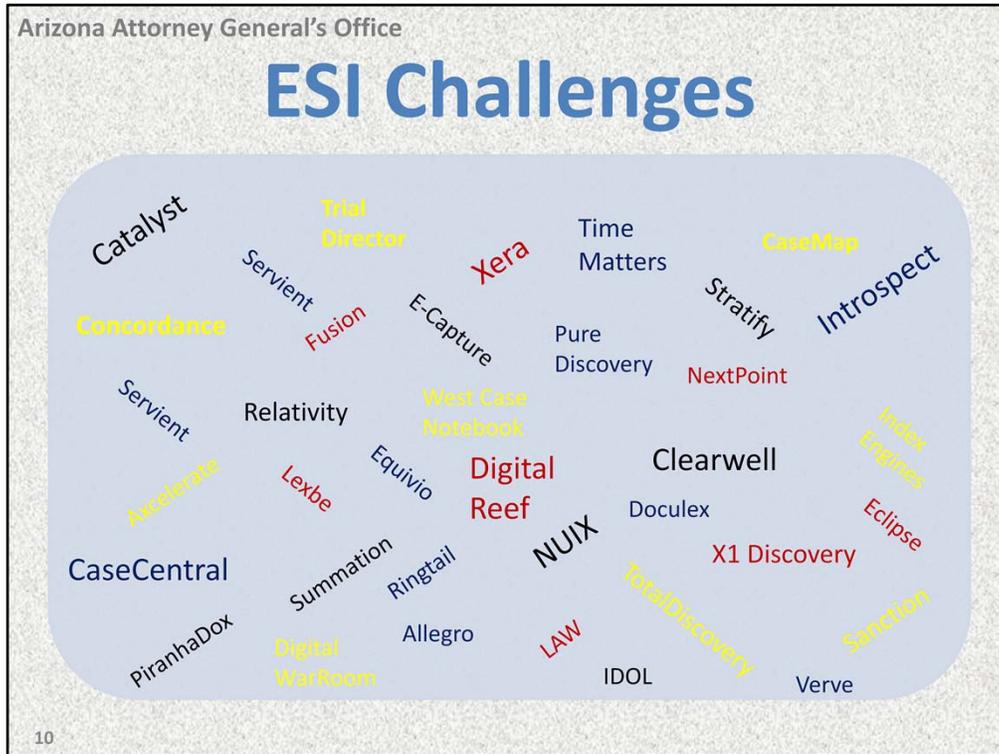
- Volume, Volume, Volume (recall the IDC Study for 2014 = 1.8 Zettabytes)
  - Location of Data
    - PC local machines, Network Servers
    - Web (internet), External Drives
    - Phones, Tablets, PDAs
    - Game Consoles, TVs, Copiers, GPS
- 
- 
-

## ESI Challenges

- **Type of Files** (aka MIME types, a term often used by vendors)  
Executables, Compressed Archives, Documents, Audio/Video, Databases, Text, Images, Etc.
- **Authenticity and Handling**  
Depending on circumstances, MD5 or SHA hash values may be required to authenticate the files

9

- The variety of file types that exist is daunting (refer back to slide #3)
  - Type of Files (aka MIME types, a term often used by third-party data processing vendors)  
Executables, Compressed Archives, Documents, Audio/Video, Databases, Text, Images, Etc.
  - Authenticity  
Depending on circumstances, MD5 or SHA hash values may be required to authenticate the files – But are they truly unique?
- 
- 
-



Dozens, if not hundreds, of applications dedicated to the handling of ESI

Collection, Processing, Review, Analysis, Production, Presentation

---

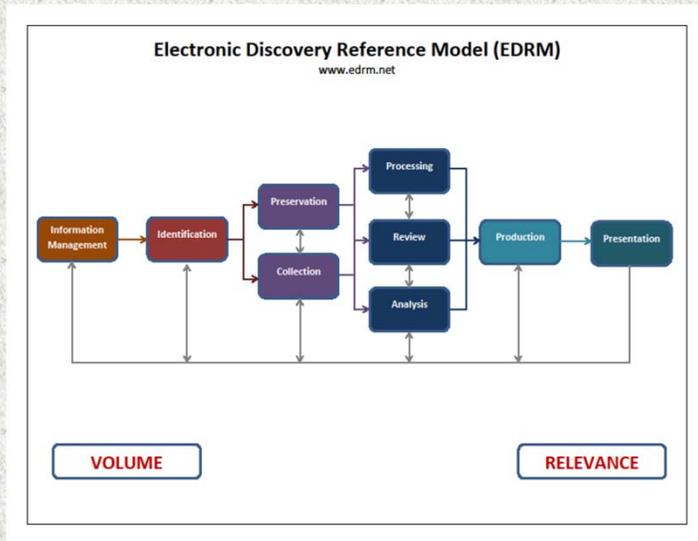


---



---

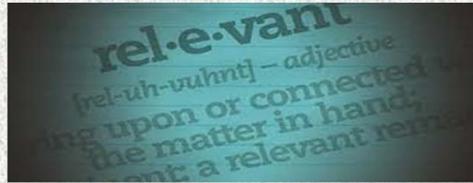
# ESI Challenges



11

- Litigation Software Tools (EDRM)
  - Collection (Forensics)
  - Processing (Data Extraction/Conversion)
  - Review (Searching / Tagging)
  - Analyzing (Strategy)
  - Producing (Internal and Opposing)
  - Presenting (Trial, Hearings)

# How will the data be used



STATE OF HAWAII		CERTIFICATE OF LIVE BIRTH		DEPARTMENT OF HEALTH	
1. Child's First Name (Type or print)		2. Middle Name		3. Last Name	
ALAN					
4. Sex		5. Date of Birth		6. Month	
Male		10/20/63		1963	
7. Place of Birth (City, Town or Rural Location)		8. Date of Birth (Day, Month, Year)		9. Time of Birth	
Honolulu		10/20/63		0:47 A.M.	
10. Name of Hospital or Institution (If not in hospital or institution, give street address)		11. Place of Birth (City or Town, County)		12. Name of Physician	
U. S. Army Tripler General Hospital		Honolulu, Hawaii			
13. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		14. Name of Father		15. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
Mehiana		Honolulu, Hawaii		Mehiana	
16. Mother's Address		17. Is Mother on a Plan or Planation?		18. Name of Father	
[Redacted]		No		[Redacted]	
19. Full Name of Father		20. Name of Father		21. Kind of Business or Industry	
[Redacted]		[Redacted]		Civilian	
22. Age of Family (1. Birthdate (month, day or Approx. Date); 2. Usual Occupation)		23. Kind of Business or Industry		24. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		Civilian		Mehiana	
25. Full Name of Mother		26. Name of Mother		27. Kind of Business or Industry	
[Redacted]		[Redacted]		Civilian	
28. Age of Mother (1. Birthdate (month, day or Approx. Date); 2. Type of Occupation (Outside Home Working Frequency))		29. Date Last Worked		30. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		10 Apr 63		[Redacted]	
31. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		32. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		33. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
34. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		35. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		36. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
37. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		38. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		39. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
40. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		41. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		42. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
43. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		44. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		45. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
46. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		47. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		48. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
49. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		50. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		51. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
52. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		53. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		54. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
55. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		56. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		57. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
58. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		59. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		60. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
61. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		62. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		63. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
64. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		65. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		66. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
67. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		68. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		69. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
70. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		71. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		72. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
73. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		74. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		75. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
76. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		77. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		78. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
79. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		80. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		81. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
82. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		83. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		84. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
85. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		86. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		87. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
88. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		89. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		90. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
91. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		92. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		93. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
94. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		95. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		96. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
97. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		98. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		99. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	
100. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		101. Name of Mother or Mother's Caregiver (Print or Initial Lastname)		102. Name of Mother or Mother's Caregiver (Print or Initial Lastname)	
[Redacted]		[Redacted]		[Redacted]	



- A Clear Purpose will Assist with ESI Handling
- Relevance and/or Privilege Review
- Fact Gathering
- Cast of Characters and Chronologies
- Redactions
- Data Productions
- Other

## How will the data be used

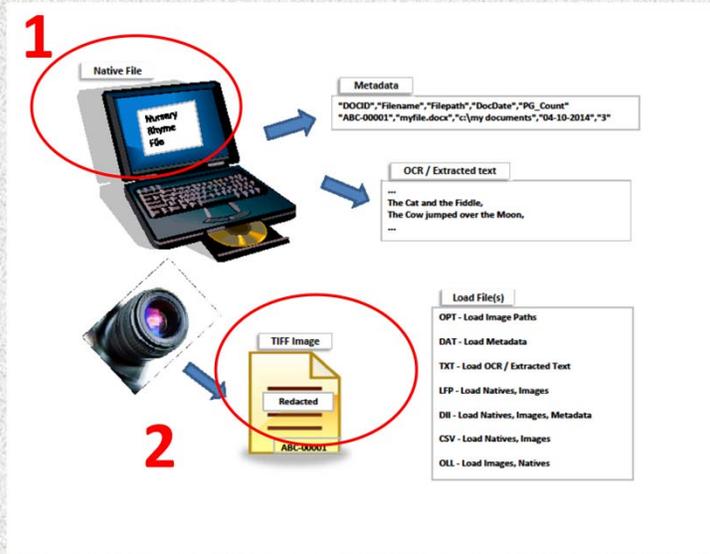
- Internal strategy & planning meetings
- FRCP 16(b)(3)(B)(ii-iii)  
Timing – Pretrial Conference Scheduling
- FRCP 26(b)  
Discovery of ESI – Scope and Limits
- FRCP 34(b)(1)(c), 34(b)(2)(E)(ii)  
Form(s) in which ESI is to be Produced

13

Keep in mind – Quantity/Volume of ESI during:

- Internal strategy & planning meetings
  - FRCP 16(b)(3)(B)(ii-iii)  
Timing – Pretrial Conference Scheduling
  - FRCP 26(b)  
Discovery of ESI – Scope and Limits
  - FRCP 34(b)(1)(c), 34(b)(2)(E)(ii)  
Form(s) in which ESI is to be Produced
- 
- 
-

# 5 'Parts' of ESI



There are 5 things you need to understand regarding ESI to effectively control Discovery data

- Natives  
Any file in the form in which it was originally created (Word DOC, Adobe PDF, Outlook MSG, Excel XLSX, Audio WAV, Video VOB, etc.)
  - Images  
An electronic picture of the document as it would look if it were photographed with a camera (BMP, TIF, JPG, PNG, GIF, etc.)
- 
- 
-

# ESI – Closer Look

The screenshot shows a file explorer window with a list of folders and one document file. The folders are named with a sequence of AZ-AGO-000000001 through AZ-AGO-00000244. The document file is named Native\_V\_Images.DOC.

Name	Date modified	Type	Size
AZ-AGO-000000001	4/15/2014 9:09 AM	File folder	
AZ-AGO-000000081	4/15/2014 9:11 AM	File folder	
AZ-AGO-000000095	4/15/2014 9:11 AM	File folder	
AZ-AGO-000000131	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000134	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000135	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000137	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000140	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000144	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000146	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000152	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000153	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000154	4/15/2014 9:12 AM	File folder	
AZ-AGO-000000240	4/15/2014 9:13 AM	File folder	
AZ-AGO-000000242	4/15/2014 9:13 AM	File folder	
AZ-AGO-000000243	4/15/2014 9:13 AM	File folder	
AZ-AGO-000000244	4/15/2014 9:13 AM	File folder	
Native_V_Images.DOC	8/14/2012 10:51 AM	Microsoft Word 9...	104 KB

- Native Files  
Any file still in the form in which it was originally created (DOCX, XLSX, PST, PPT, MPG, HTML, etc.)
- Contains relevant metadata
- Typically requires the original application to review
- Typically cannot be Bates Stamped or Redacted
- Less expensive to acquire when compared to Full Image Processing

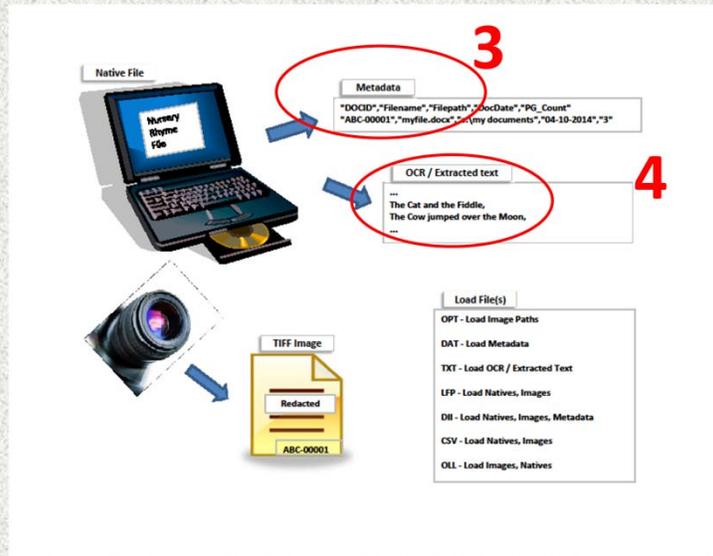
# ESI – Closer Look

Name	Date modified	Type	Size
AZ-AGO-00000001.jpg	4/15/2014 9:09 AM	JPEG image	261 KB
AZ-AGO-00000002.jpg	4/15/2014 9:09 AM	JPEG image	541 KB
AZ-AGO-00000003.jpg	4/15/2014 9:09 AM	JPEG image	656 KB
AZ-AGO-00000004.jpg	4/15/2014 9:09 AM	JPEG image	605 KB
AZ-AGO-00000005.jpg	4/15/2014 9:09 AM	JPEG image	775 KB
AZ-AGO-00000006.jpg	4/15/2014 9:09 AM	JPEG image	771 KB
AZ-AGO-00000007.jpg	4/15/2014 9:09 AM	JPEG image	826 KB
AZ-AGO-00000008.jpg	4/15/2014 9:09 AM	JPEG image	879 KB
AZ-AGO-00000009.jpg	4/15/2014 9:09 AM	JPEG image	697 KB
AZ-AGO-00000010.jpg	4/15/2014 9:09 AM	JPEG image	690 KB
AZ-AGO-00000011.jpg	4/15/2014 9:09 AM	JPEG image	783 KB
AZ-AGO-00000012.jpg	4/15/2014 9:09 AM	JPEG image	384 KB
AZ-AGO-00000013.jpg	4/15/2014 9:09 AM	JPEG image	339 KB
AZ-AGO-00000014.jpg	4/15/2014 9:09 AM	JPEG image	371 KB
AZ-AGO-00000015.jpg	4/15/2014 9:09 AM	JPEG image	380 KB
AZ-AGO-00000016.jpg	4/15/2014 9:09 AM	JPEG image	470 KB
AZ-AGO-00000017.jpg	4/15/2014 9:09 AM	JPEG image	751 KB
AZ-AGO-00000018.jpg	4/15/2014 9:09 AM	JPEG image	222 KB

16

- Image Files  
Any file converted from its Native format into an image file (a picture)  
(JPG, BMP, TIF, PDF)
- No longer contains original metadata
- Can be viewed by common Image Viewers
- Can be annotated (Bates Numbers, Redactions)
- Significantly more expensive than Natives

# 5 'Parts' of ESI



17

- Metadata  
Background information about file typically not readily available to the average user (file statistics, creation date, system author, modification date, document versions, etc.)
- Extracted Text / OCR  
Machine readable/searchable characters – The 'text' you see in a document may, or may not, be readable by the computer

# ESI – Closer Look

```
pBEGDOCp pENDDOCp pPAGECNTp pAccess_Datep pCreate_Datep pApplication_Namep pAttachment_Countp pAttachment_Li
pAZ-AGO-00000001p pAZ-AGO-00000080p p80p p11/12/2012 9:27:50 FMP;:p pMicrosoft Word 2010p p p p p p p p
pAZ-AGO-00000081p pAZ-AGO-00000094p p14p p07/29/2011 7:44:41 FMP;:p pMac OS X 10.6.8 Quartz PDFContextp p
pAZ-AGO-00000095p pAZ-AGO-00000130p p36p p08/22/2012 1:35:21 FMP;:p pAdobe PDF Library 10.0p p p p p p p p
pAZ-AGO-00000131p pAZ-AGO-00000133p p3p p10/24/2013 10:41:00 AMP;:p pMicrosoft Office Wordp p p p p p p p
pAZ-AGO-00000134p pAZ-AGO-00000134p p1p p6/10/2013 7:07:31 AMP;:p pMicrosoft Excelp p p p p p p p p p
pAZ-AGO-00000135p pAZ-AGO-00000136p p2p p8/16/2012 9:06:03 AMP;:p pMicrosoft Excelp p p p p p p p p p
pAZ-AGO-00000137p pAZ-AGO-00000139p p3p p8/15/2012 4:23:00 FMP;:p pMicrosoft Office Wordp p p p p p p p
pAZ-AGO-00000140p pAZ-AGO-00000143p p4p p8/20/2012 1:48:05 FMP;:p pMicrosoft Excelp p p p p p p p p p
pAZ-AGO-00000144p pAZ-AGO-00000145p p2p p8/14/2012 10:51:00 AMP;:p pMicrosoft Office Wordp p p p p p p p
pAZ-AGO-00000146p pAZ-AGO-00000151p p6p p4/26/2013 2:39:20 FMP;:p pMicrosoft Excelp p p p p p p p p p
pAZ-AGO-00000152p pAZ-AGO-00000152p p1p p1/31/2013 12:30:38 FMP;:p pMicrosoft Excelp p p p p p p p p p
pAZ-AGO-00000153p pAZ-AGO-00000153p p1p p9/11/2012 9:37:00 AMP;:p pMicrosoft Office Wordp p p p p p p p
pAZ-AGO-00000154p pAZ-AGO-00000239p p86p p03/15/2006 2:03:00 FMP;:p pCorel PDF Engine Version 11.2.0.239p
pAZ-AGO-00000240p pAZ-AGO-00000241p p2p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000242p pAZ-AGO-00000242p p1p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000243p pAZ-AGO-00000243p p1p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000244p pAZ-AGO-00000245p p2p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000246p pAZ-AGO-00000249p p4p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000250p pAZ-AGO-00000251p p2p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000252p pAZ-AGO-00000253p p2p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000254p pAZ-AGO-00000254p p1p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000255p pAZ-AGO-00000255p p1p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
pAZ-AGO-00000256p pAZ-AGO-00000256p p1p p;:p;:p p p p p p p p p p p p p p p p p p p p p p p p p p p p
```

- Metadata  
Information about the original file that is embedded within the file and often not easily accessible by the average user
- Dates: Creation, Modified, Saved, etc.
- File Type, Author, image statistics, etc.
- Revisions (depending on application)

## ESI – Closer Look

- Extracted Text / OCR
  - Extracted text is the actual text as it was typed into the application that created the file (Word, Excel, Email, etc.)
  - OCR (Optical Character Recognition) is the computer's attempt at converting a picture into machine readable text (TIF, PDF, JPG, etc.)

19

- Extracted Text / OCR
    - Extracted text is the actual text as it was typed into the application that created the file (Word, Excel, Email, etc.)
    - OCR (Optical Character Recognition) is the computer's attempt at converting a picture into machine readable text (TIF, PDF, JPG, etc.)
- 
- 
-

# ESI – Closer Look

## Good Image – Excellent OCR Quality

IMAGE

### Litigation Support Manager – Key to a Successful Case

We've all heard about the changes to the ABA rules. Specifically Rule 1.1 as it relates to Competence. The proposed change states "To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology..."

TEXT

### Litigation Support Manager – Key to a Successful Case

We've all heard about the changes to the ABA rules. Specifically Rule 1.1 as it relates to Competence. The proposed change states "To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology..."

## Good Image – Excellent OCR Quality

Original document converted to high quality black & white image



# ESI – Closer Look

## Poor Image – Poor OCR Quality

IMAGE

### Litigation Support Manager – Key to a Successful Case

We've all heard about the changes to the ABA rules. Specifically Rule 1.1 as it relates to Competence. The proposed change states "To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology..."

TEXT

### Litigation Support Manager – Key to a Successful Case

We've all heard about the changes to the ABA rules. Specifically Rule 1.1 as it relates to Competence. The proposed change states "To maintain the requisite knowledge and skill, a lawyer should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology..."

## Poor Image – Poor OCR Quality

Original document converted color/grayscale image  
Notice the background 'noise' plays havoc with the text



## ESI – Closer Look

Something to Consider:  
On any given page, the average word length  
is 5 characters.

If you receive OCR with 80% accuracy  
(1 out of every 5 characters are errors)

Nearly every word could be an error (29)  
Noarly eUery worj aould be an emor (6)

22

Something to Consider:

On any given page, the average word length is 5 characters.

If you receive OCR with 80% accuracy  
(1 out of every 5 characters are errors)

Nearly every word could be an error (29)  
Noarly eUery worj aould be an emor (6)

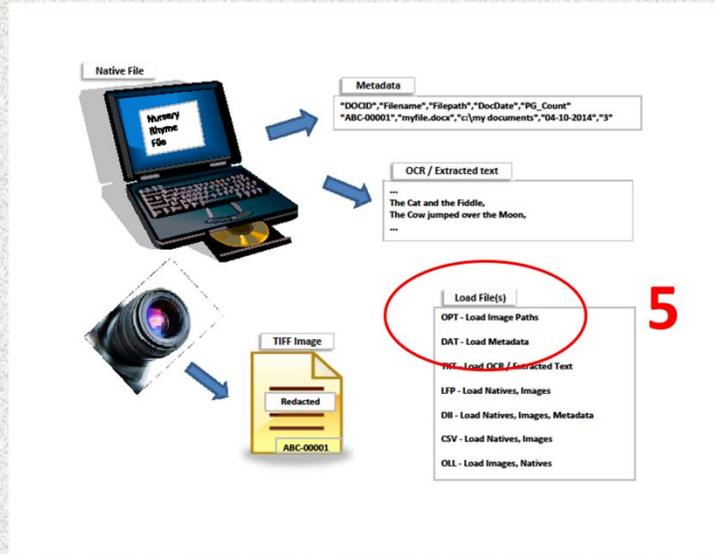
For searching, high quality OCR is paramount

---

---

---

# 5 'Parts' of ESI



## Load Files

Although not technically part of any specific document/file, Load Files are most often just formatted text files and are necessary for loading ESI into the tools/applications we will ultimately be using (DAT, OPT, DII, TXT, CSV, LFP, etc.)

Load files are incredibly important especially when delivering Images since they supply document break information and it is the only way to supply the associated metadata

# ESI – Closer Look

Name	Date modified	Type	Size
IMG_0000001	4/15/2014 10:26 AM	File folder	
OF_0000001	4/15/2014 9:13 AM	File folder	
TXT_0000001	4/15/2014 9:13 AM	File folder	
CONCORD.DAT	4/15/2014 9:13 AM	DAT File	8 KB
Exportlog.log	4/15/2014 9:13 AM	LOG File	23 KB
IPRO.LFP	4/15/2014 9:13 AM	LFP File	40 KB
IPROCOR.LFP	4/15/2014 9:09 AM	LFP File	0 KB
OFXRef.txt	4/15/2014 9:13 AM	TXT File	3 KB
OPTICON.OPT	4/15/2014 9:13 AM	OPT File	26 KB
PROJECT.SFP	4/15/2014 9:13 AM	SFP File	1 KB
SUMMAT.DII	4/15/2014 9:13 AM	DII File	20 KB

24

## Load Files

Even though Load Files are not actually related to a specific file, they are the means by which we can load the ESI into the various Litigation Support applications used to Review, Analyze, and Present the files.

- Load Files are most often simple delimited text files, although they can be complex text files and sometimes even binary data files
  - The most common extensions for these files are: TXT, CSV, OPT, LFP, DII, DAT
- 
- 
-

## ESI – Closer Look

- Load Files should contain a unique record ID or Bates Number
- Commonly contains document boundary information – Without this information, it is nearly impossible to effectively review and produce relevant material

25

- Load Files should contain a unique record ID or Bates Number
- Commonly contains document boundary information – Without this information, it is nearly impossible to effectively review and produce relevant material

---

---

---

## ESI – Closer Look

- Load Files may also contain path information to the original Native files and/or the converted Image files
- Depending on your requirements, the Load Files may also contain Metadata and sometimes even the Extracted or OCR text (although Text is normally produced as separate files and not included in the Load Files)

26

- Load Files may also contain path information to the original Native files and/or the converted Image files
- Depending on your requirements, the Load Files may also contain Metadata and sometimes even the Extracted or OCR text (although Text is normally produced as separate files and not included in the Load Files)

---

---

---

# Standard Request

- **Native Files Request**
- Native Files as a forensic image  
**OR**
- Native Files with original Metadata  
**and**
- OCR/Extracted Text if available

27

Least expensive and typically the quickest Form of Production  
May require additional processing depending on usage  
Requires appropriate applications to review documents

---

---

---

# Standard Request

- **Image Files Request**

- CCITT Group IV Tiff files (JPG for color images)
- Metadata (Concordance delimiters 20/254)
- OCR (1 text file per document)
- Image Load File (OPT or LFP w/doc designations)
- Unique Sequential DocID's
- Filenames should match Bates Numbers (DocID's)
- 300 DPI minimum

28

Most expensive Form of Production

Very large amount of data compared to a Native request

Takes longer to produce than a Native request

Does allow for Redactions

Can be Bates stamped or Annotated

---

---

---

# Standard Request

- **PDF Files Request**

- PDF/A format
  - With embedded searchable text or,
  - OCR text files (1 text file per document)
- 1 PDF file per original document
- Metadata (Concordance delimiters 20/254)
- Filenames should either
  - Match original Native filenames or,
  - Unique sequential Bates Numbers (DocID's)

29

Expensive (like an Image request)

Takes longer to produce than a Native request

Often requested when the document population is small and a database will not be used for review

Can also be Redacted or Annotated

---

---

---

# Example Scenario



We've covered the items we may need and why we need them, now, to make more sense of it all, we will tie it all together in a real-world scenario

This example revolves around our current Litigation Support products: Concordance, CaseMap, and Trial Director

---

---

---

## Example Scenario

- Little Business (LB) v Mega Corporation (MC)
- Deceitful Marketing Practices
- LB has a single hard drive of data
- MC has dozens of custodians and data locations
- Due to available funds and the minimal potential award, only 3 of the most critical MC drives and the LB drive will be processed and searched

31

Proportionality is a key consideration when deciding what data to process

Timing, as in any case, must also be considered when processing ESI

Based on our research, we've determined that there are only 3 key drives at MC

Combined with our LB drive, we should be able to locate all the information we need for our case

---

---

---

## Example Scenario

- **Our Production to Opposing (MC):**
    - Need to locate and redact privileged information
    - We have a small/limited number of files
    - We will process files and produce Images with Metadata and OCR
- MC agrees to this in our 26(f) conference

32

Our data must be reviewed for privilege (and possibly redacted) before sending to opposing counsel for MC

In our 26(f) conference, MC counsel agrees to accept images and metadata

---

---

---

## Example Scenario

- **What we will do with the LB files:**
  - Process files to generate images and to extract metadata and searchable text
  - Load images, metadata, and searchable text into Concordance to search for and redact privileged information
  - Export images, metadata, and OCR for production to MC (care must be taken to exclude redacted information from OCR and metadata)

33

The processing software will create TIFF/JPG images that can be redacted and it will extract the text we need to load to allow us to search the documents

It will also extract the metadata that we will need to include in our production

---

---

---

## Example Scenario

- **Our request to Opposing (MC):**
  - Need to review for relevance
  - Need to organize for case strategy
  - Need to produce electronically in court
  - Potential for hundreds of thousands of files
- Will request Native Files and Metadata (we will process to generate searchable text)  
MC agrees to this in our 26(f) conference

34

We want to locate relevant documents to prepare our case and we will produce them electronically at trial

There may be hundreds of thousands of documents to review

Based on the above information, we determine that we need native files because there may be a large number of databases and spreadsheets that would not be useful in image format

Counsel for MC agrees to provide native files and metadata (we need metadata since these will not be forensic drive images)

---

---

---

## Example Scenario

- **What we will do with the data (LB and MC):**



- Load native files and searchable text into Concordance to search for and locate relevant records (Tag results)



- We will load all relevant documents into CaseMap which will allow us to analyze the facts and other case information



- Create electronic exhibits from key documents and load them into Trial Director for courtroom presentation

Concordance helps us find the relevant documents  
CaseMap allows us to analyze the facts and strategize  
Trial Director presents our exhibits electronically

---

---

---

## Sample Production Letter

We are producing documents electronically as single-page TIFF files titled as the Bates number, document level text files containing extracted text and OCR text, a load file containing metadata, and an Opticon image cross-reference file.

We are also including in this production links to the native video files for certain documents.

Additionally, we are also including in this production links to the native Excel spreadsheets for certain documents.

Please note that, as with documents we have previously produced in native format (including Excel spreadsheets), the native files do not contain Bates labels or confidentiality designations.

Each native file shall be construed as containing the same Bates number and confidentiality designation (if any) as its corresponding TIFF image file.

This is an actual Production Letter from an active case

The various 'parts' of ESI have been mixed into a hybrid production



## Additional Considerations

- Costs
- Unitization Nightmares
- Bates Numbers vs Image Keys vs File Names
- Media and/or Data Transmission
- Single vs Rolling Deliverables

37

- Costs based on volume and format (and timing)
  - Unitization Nightmares – Single documents merged into one file or split into multiple files
  - Bates Numbers vs Image Keys vs File Names (paper/image, database, computer file)
  - Media and/or Data Transmission (security, volume, timing)
  - Single vs Rolling Deliverables
- 
- 
-

## Additional Considerations

- Data Collection  
Location / Media Type(s)
- Internal vs Remote Resources for Review
- FRCP 26(f) Meet and Confer Conference Stipulations

38

Data Collection – Location and/or Media Type(s)

Internal vs Remote Resources for Review

FRCP 26(f) Meet and Confer Conference Stipulations

- Form(s) of Production
- Search Terms / Protocol (see next slide)
- Proportionality (cost/effort vs. potential reward)
- Timing of Discovery and Production(s)
- Custodians (do they have relevant ESI)
- Data (Reasonably Accessible, Volume, Data Type(s))
- Privileged Document Handling
- Potential for Data Sampling

Ask for help BEFORE the Conference

---

---

---

## Additional Considerations

Some additional review considerations:

- Predictive Coding
- Near-Dupe Determination
- De-NISTing
- Email Threading
- Data Analytics, Latent Semantic Indexing, Term Necessity Prediction

39

### Near-Dupe Determination

(Not an exact duplicate – only a set % is identical)

### De-NISTing (National Institute of Standards and Technology)

(List of standard files – often excluded from productions)

### Email Threading

(Email conversation strings – limits duplicative email review)

### Predictive Coding, Seed Sets, Data Analytics, Latent Semantic Indexing, Term Necessity Prediction

(These terms are related to TAR – Technology Assisted Review)

(Often used by opposition to test the competition)

---

---

---

# Final Consideration



40

- FRCP Rule 26(g)
  - Signature from the Attorney of Record
  - Certifies that Disclosures are “complete and correct” and,
  - A Discovery Request is “consistent with these rules...” and is “neither unreasonable nor unduly burdensome or expensive...”
  - FRCP 26(g)(3) Sanction for Improper Certification  
Places the burden on the signer to understand the technology behind ESI
- 
- 
-

## Conclusion

- This CLE is just a sample of the technology issues that you will face in modern litigation
- It may have been a refresher for some of you, but it was likely an eye-opener for many
- The primary intent of this CLE session is to, hopefully, generate more questions than answers to get the mental wheels turning regarding ESI

41

- This CLE is just a sample of the technology issues that you will face in the modern litigation
  - It may have been a refresher for some of you, but it was likely an eye-opener for many
  - The primary intent of this CLE session is to, hopefully, generate more questions than answers to get the mental wheels turning regarding ESI
- 
- 
-

Feedback

Please feel free to contact me if you have any questions or would like more information



# Feedback

---

---

---