Developing a Defensible eDiscovery / Forensics Program for any Organization

October 22, 2014
Start Time: 11:30 am US Eastern
ISSA International Conference
Orlando, Florida, USA

J. Michael Butler
GCFA EnCE GSEC CISA

ISSA
Information Systems Security Association
Disclaimer

The opinions expressed herein are independent views of the author and do not necessarily reflect the views and opinions of the author’s current employer or any other entity.

Author retains full rights.
eDiscovery and Computer forensic capabilities are crucial in our increasingly litigious society.

I recently queried three different firms who offer computer forensic and eDiscovery services as outside consultants. I was looking for a commonly accepted hourly price for an outside firm to come in to assist with forensics and eDiscovery. The average of the three firms was $265 per hour, but the price can go up markedly if they are asked to provide expedited services or appear in court as experts. The costs really go through the roof if your matter requires outside review. According to a 2012 article in “Inside Counsel” magazine, costs for review are estimated anywhere between $5,000-$30,000 per gigabyte. They quote a Rand survey who put the going rate at $18,000 per gigabyte.

The decision each organization has to make is whether they want to outsource this function, or control eDiscovery and computer forensics internally. One major case can more than justify the expenditure for software tools and personnel required for internal computer forensics and eDiscovery capabilities. The possible sanctions associated with NOT being able to produce what the court has specified can far outstrip the cost of archiving systems and eDiscovery software and personnel.

We will discuss options for archiving and retrieving email, accessing and protecting data properly from user's devices, as well as from company systems of record, and

---

**Are You Ready?**

- Is your organization prepared to produce data demanded by the court during litigation?
- Are you in a position to support your Legal Department’s efforts to defend your organization?
- What about the costs of eDiscovery?
- “Forensic science is generally defined as the application of science to the law.” – NIST

- [http://www.insidecounsel.com/2012/05/23/e-discovery-costs-pay-now-or-pay-later](http://www.insidecounsel.com/2012/05/23/e-discovery-costs-pay-now-or-pay-later)
the plethora of other locations where data live. From a practitioner’s viewpoint, determine the steps that you must take to move your organization into a defensible position **before** opposing counsel makes their requests and challenges your practices.
Today we are discussing two very closely related but different topics. Forensics is the basic idea of collecting, analyzing, and reporting on data recovered in an investigation. eDiscovery is electronic Discovery as it relates to litigation. eDiscovery is usually negotiated between the two parties in a suit, where they will decide together who is going to provide what data for the matter. eDiscovery is the same process, basically, as any trial “Discovery” phase except the “e” refers to Electronically Stored Information or “ESI” as opposed to physical evidence. eDiscovery certainly includes digital forensics and the forensics analysis required in order to provide the ESI requested for a matter.

eDiscovery started out as a process, but has become an entire marketplace where vendors can provide software or consulting or both to assist an organization with discovery of their ESI for litigation.

Let’s start our discussion today with some of the basic requirements of a solid Computer Forensics program for any organization. Then we will see how that dovetails in with the eDiscovery processes.
There are a number of excellent books and online resources. Here are two.

For the technical analyst, the classic is Brian Carrier’s *File System Forensic Analysis*. It is brand agnostic and goes way down into the weeds in regard to file systems and how they work.

A new book by Suzanne Widup, called *Computer Forensics & Digital Investigation with EnCase Forensic v7*, (which would be a great addition for anyone with an EnCase shop), has as chapter one “The Road to Readiness” which really has little to do with EnCase and everything to do with preparation for ANY kind of forensics in your organization. She said...
Forensic Readiness

- “...the state of an organization that has completed the effort to properly prepare for performing digital forensic investigations. Policies are supported by a foundation of repeatable processes and trained, well-equipped personnel.”

- Do you have a digital forensics policy?

- What processes does the policy point to?

The number one requirement for being ready is having a Policy or including in your Information Security policies statements that support the “...collection, preservation, and analysis of digital evidence for eventual presentation in court.” [ibid] Only AFTER you have a policy in place are you ready to acquire a facility, equipment, and other resources that you need for forensic collection and analysis, and establish procedures for that work.

While it isn’t our goal to write a policy today, let’s look at some basics you will want to address...
First comes the policy, then comes the process. Policy must address how an investigation can be initiated, by whom, and who can then approve the investigation IN WRITING, all before the process can start. I carry a card with me [get it out] that is my get out of jail free card. The really smart techs at your organization, who would be thrilled to be asked to go find information on someone thus showing off their incredible skills, must be reigned in. By using normal administrative techniques to search for data or copy it somewhere else for analysis, a technician without appropriate training will stomp all over evidence and make it unusable in court.

Who is qualified in your organization to do Forensic Analysis? What qualifies them? Being really technically smart is not enough. The investigator MUST understand the legal requirements and ramifications of what he or she will do during an investigation. So training is extremely important, and I would recommend certification as well, if it is a reputable cert that really means the analyst knows what he or she is doing. There are several certifications for Computer Forensics. EnCE is specific to EnCase used by most law enforcement. GCFA is a GIAC cert. CCFP is an ISC² certificate. The CCE is sanctioned by the The International Society of Forensic Computer Examiners or ISFCE. There are other certs as well, but these may be the more well known.

In a nutshell, such training will teach your analyst that all collections, examinations, and analysis must be performed in a Forensically Sound Manner. This is a phrase you will want to take to heart, especially when talking about matters with Management
and Legal. Let us take a look at the Forensic Processes to be performed.
There are four basic steps in the Forensic Process as defined by NIST. Your organization must be involved at some level in all of these steps, but none of them are possible without the correct preparation. Notice in this cyclical representation of the process, from the NIST SP800-86 document, that one can expect to cycle back to the beginning and run through the process multiple times. It would be an even better representation if the arrows would lead back to the beginning from any and all of the 3 advanced stages. Examination can lead to a determination that something else must be collected. So can analysis. At any point during processing, some new bit may come to light that would point us back to the Collection phase in order to search for previously unknown relevant data.

When a third party forensics firm is involved in this process, they add layers of complexity and create more work, more resource demands, and more expense, because they don’t know your business like you do.

The question is, what can your organization do to prepare for the time you will be called on to perform computer forensics and/or eDiscovery in order to appropriately respond to a legal matter? Are you prepared to react at either a Physical or a Logical level?...
Logical collections refer to the data being collected, not the media.

Physical collection refers to physically acquiring, tracking, and storing physical devices that could contain data. Your policy and processes must address how you go about such collections. Such collections may require that you be accompanied by someone from HR or Security when you go to acquire devices. In any case, you should be prepared to track what you collect using a good Chain of Custody form so that all details about the possession of the item collected are tracked on paper or with software in such a way that you can provide those details at a later date.

Create a Chain of Custody form that has spaces for writing all this information. I have included a link to a standard NIST CoC form, but if you simply Google: “Chain of Custody form,” you will find many examples of forms that you can adapt for your use. Work closely with Legal to make sure they are happy with your form, because they are the folks that have to defend it. And if they can’t, you probably will get to.

Whether you handle your collections in house or you hire a third party, the mantra will remain the same. All collections must be performed in a “Forensically Sound Manner”...

Let’s look at a list of what to record on a Chain of Custody form...
Details like who collected the item, when did they collect it, what identifying marks are on the item – such as asset tags, serial numbers, any marks the analyst added, etc. Where was it collected? Where was it stored? Who else had possession of the item, when, why, and for how long? How was it kept secure?

Chain of Custody information can also be tracked electronically, though you need to give consideration to how easy it would be to manipulate the electronic data, and what steps you might take to ensure that it is not altered once entered. A properly configured database might be a good way to track such data.

To help validate the CoC, use date and time stamped photography of the items being collected. Store that information with your case data and use hashes to verify the files have not changed.

Collection is the first step of the forensic analysis process...
No matter what tool or method you use to collect and preserve your data, you must ensure that you collect both the metadata and the data without altering one bit of the data. Again, “In a Forensically Sound Manner.”...
Data is collected in one of two ways. Historically, forensic processes have always involved collection of physical data – blood, DNA, bullets, and the bodies. Oh, yeah, and hard drives, thumb drives, smart phones, printers, cameras, memory cards – in short any physical device that can contain data. Now from those physical devices, we also collect logical data – files, lost files, deleted files, files and file fragments in unallocated space, etc.

What does “Forensically Sound” mean? To show that your data is forensically sound, you need to be able to prove that the data have not changed since they were collected, including the metadata. (We will have further discussion about metadata in a few minutes.)

Who has had the data since you collected it, and what have they done with it? Can you prove that neither you nor they did anything to alter the data? One important concept to understand is the use of hashing algorithms to prove the original data has not changed.

Every piece of information collected also is accompanied by “metadata.” So what is metadata?...
The “Duh” definition is the one that everyone throws back at you – metadata is “data about data.” While technically correct, this is not helpful. Reminds me of the old saw about the hot air ballooner who was floating about on a cloudy day with very limited visibility, so he pulls the rope and descends to about 50 feet off the ground, spies a man on the ground, and shouts out to him, “where am I?” To which the man replies, “You are in a balloon about 50 feet off the ground!” The balloonist says, “Thank you, counselor.” The man on the ground says, “How did you know I was a lawyer?” “Because,” the balloonist responds, “Your answer is technically and defensibly accurate and totally useless.”

I will qualify that story with the statement that none of the lawyers I work with ever say anything that is useless. In case they are here... listening...

While metadata is most assuredly information about data, a more useful definition of metadata would tell us that metadata describes the history, including tracking or management, of an electronic document. Metadata can be located internal to the document. For example...

The metadata definition is the one that everyone throws back at you – metadata is “data about data.” While technically correct, this is not helpful. Reminds me of the old saw about the hot air ballooner who was floating about on a cloudy day with very limited visibility, so he pulls the rope and descends to about 50 feet off the ground, spies a man on the ground, and shouts out to him, “where am I?” To which the man replies, “You are in a balloon about 50 feet off the ground!” The balloonist says, “Thank you, counselor.” The man on the ground says, “How did you know I was a lawyer?” “Because,” the balloonist responds, “Your answer is technically and defensibly accurate and totally useless.”

I will qualify that story with the statement that none of the lawyers I work with ever say anything that is useless. In case they are here... listening...

While metadata is most assuredly information about data, a more useful definition of metadata would tell us that metadata describes the history, including tracking or management, of an electronic document. Metadata can be located internal to the document. For example...

The metadata definition is the one that everyone throws back at you – metadata is “data about data.” While technically correct, this is not helpful. Reminds me of the old saw about the hot air ballooner who was floating about on a cloudy day with very limited visibility, so he pulls the rope and descends to about 50 feet off the ground, spies a man on the ground, and shouts out to him, “where am I?” To which the man replies, “You are in a balloon about 50 feet off the ground!” The balloonist says, “Thank you, counselor.” The man on the ground says, “How did you know I was a lawyer?” “Because,” the balloonist responds, “Your answer is technically and defensibly accurate and totally useless.”

I will qualify that story with the statement that none of the lawyers I work with ever say anything that is useless. In case they are here... listening...

While metadata is most assuredly information about data, a more useful definition of metadata would tell us that metadata describes the history, including tracking or management, of an electronic document. Metadata can be located internal to the document. For example...

The metadata definition is the one that everyone throws back at you – metadata is “data about data.” While technically correct, this is not helpful. Reminds me of the old saw about the hot air ballooner who was floating about on a cloudy day with very limited visibility, so he pulls the rope and descends to about 50 feet off the ground, spies a man on the ground, and shouts out to him, “where am I?” To which the man replies, “You are in a balloon about 50 feet off the ground!” The balloonist says, “Thank you, counselor.” The man on the ground says, “How did you know I was a lawyer?” “Because,” the balloonist responds, “Your answer is technically and defensibly accurate and totally useless.”

I will qualify that story with the statement that none of the lawyers I work with ever say anything that is useless. In case they are here... listening...

While metadata is most assuredly information about data, a more useful definition of metadata would tell us that metadata describes the history, including tracking or management, of an electronic document. Metadata can be located internal to the document. For example...
...in the case of a JPEG picture, internal Exif or (Exchangeable Image File Format) data could include a host of information about the picture that was taken, as well as the person taking it...

**Metadata (cont.)**

- **Internal**
  - Part of the Properties Information (Exif) buried in the file itself, such as a JPEG file

  - [http://www.techterms.com/definition/metadata](http://www.techterms.com/definition/metadata)
Internal Exif Metadata

- In the properties of a file, such as Exchangeable image file format ("Exif") data regarding images that can include:
  - Date/time taken
  - GPS latitude/longitude location
  - Color balance
  - F-stop
  - Exposure time

Date, time, latitude, longitude, camera settings...
More camera settings and camera information, key words, other metadata about the file itself. So where do you find this data normally?...
If you pull up the properties of many files, the OS is smart enough to locate and display the internal file properties. Here, for example, are some properties for a picture I took for this presentation. In this case, it includes the date and time, the Camera make and model and settings, along with other details that can, by the way, be removed if you click on the circled link...
Whether by malicious intent, or to simply protect one’s private information, it is a simple matter to go into the file properties, click the URL that says “Remove Properties and Personal Information,” then go to this screen where you can check the items you want removed. Note that these screen shots came from a Windows 8 box, so your version may look slightly different. In any case, with any OS, there are applications available that will allow you to view, remove, or change file properties. That being said, most folks do not have time or take the time to change file properties, or there may be some metadata that cannot be so easily changed, so it is important to collect this metadata without changing it, and store it for future reference in any legal matter. But this is just one type of file metadata stored within one type of file. There are many metadata items stored with other file types as well. Let’s look at a few well known types...
Microsoft Office documents may contain a host of information that you don’t even know might have been saved with the document. Indeed, any application may store specific information about the history of a file within the file itself. To keep these documents “forensically sound,” means we have taken pains to ensure that not one piece of this metadata changes within the file. In fact, we must ensure that not one bit of the file has changed. Just as the medical examiner must make absolutely certain that a blood sample or collected hair is not contaminated, thus rendering the evidence unusable, the digital forensic examiner must ensure that neither data, nor metadata are altered in the process of analysis. Let’s take a quick look at some metadata associated with a Microsoft Office document...

Internal Document Metadata

- Part of the internal file structure or protocol of compound files such as email headers, Docx, Xlsx, PPTx,
- Documents can contain:
  - Author
  - Initials
  - Company Name
  - Edit date(s)
  - Sender
  - Recipient
  - Path of email
  - Much more...
Here is a screen shot of the properties of this presentation that I took as I worked on it. Note it contains the Title, author, last editor, number of revisions, create date, last saved date, and even the total time spent editing so far, as well as other properties that are not on the current screen. These are all internal metadata. What about external?...
External metadata is stored somewhere else. Like the library card tells us all about the book on a shelf somewhere else, the external metadata can give us facts about the collected data. Examples of external metadata could include the computer MFT where the file names and other file information is stored, but not always the data. Or the File Allocation Table, or FAT, where name and relevant dates are stored. Or in Unix type systems, the inode or “Index” node where file information is stored. Any log file could be another repository for metadata – or history about data – also stored separate from the data as well.

What can we learn about a file from the Windows MFT, for example?...
Possibly the most significant metadata in any matter will be this file metadata. These data can go to prove who knew what and when they knew it. Every file in the NTFS or Windows file system, for example, will be accompanied by these pieces of information. When the file was Modified, Accessed, or Created – which provides the MAC times acronym most commonly used when discussing files. We can also determine when a file was last changed, or if it has been deleted, we can determine when that happened. I have spent a good deal of time discussing Metadata because it is so crucial to any case. Therefore it is absolutely imperative that we collect metadata intact for a matter, or we must be prepared to explain why it is not accurate. This could make or break your case...
So Metadata is an important part of the overall forensic process that starts with Collection. For a Forensic Process, the next steps are to Examine, Analyze, then Report on what has been collected. These processes can be detailed in Runbooks for your organization...
The idea behind “Runbooks” is to create a “Runbook,” or a documented process, for every possible scenario. So, for example, in a case where an employee is suspected of harassing another employee, the run book might call for specific steps to take to acquire approval from Legal, collect evidence from systems, such as eMail and Instant Message, as well as a process for searching the employee’s hard drive for evidence of documents or file fragments that might prove or disprove the case.

The challenge is to think of all scenarios for which run books are needed, to write all those runbooks, then update them as they are tested in real life. This is all part of the preparation for Forensics and/or eDiscovery.

Now let’s look at the second step in the forensic analysis process...
Examining the data also includes eliminating as much irrelevant data as possible. If a hard disk has 37,000 documents on it, but only 10 are relevant, then the Examination phase is where we eliminate those that are not needed. This is the phase where we narrow down the existing collection of ESI to that which might actually have relevance to the matter at hand. In eDiscovery activities, this is kin to the “culling” activities that eliminate as much unnecessary data as possible in order to make it practical to review what is left.

For examination, an organization can use their internal forensic staff, which can be one person or a team of people, who will work together with Legal – inside or external counsel – to define exactly what data could possibly be relevant. Then the forensic staff can use appropriate tools to access the data and export out any items that might be relevant to the issue using search terms provided by counsel. Alternatively, as with any phase of Forensics or eDiscovery, a third party firm can be hired to collect, examine, analyze, and provide reporting regarding a litigious matter.

Step 3 in the Forensic Process:...
Analysis

“The foundation of forensics is using a methodical approach to reach appropriate conclusions based on the available data or determine that no conclusion can yet be drawn. The analysis should include identifying people, places, items, and events, and determining how these elements are related so that a conclusion can be reached.” – NIST SP800-86 p. 3-6

Analysis is taking the facts discovered in the digital evidence and identifying them for Legal, HR, or Management, as appropriate. It is important to note that, while the NIST definition calls for drawing conclusions, it does not mean that the analyst should be relying upon his or her opinion. The best forensic analysis will provide “just the facts,” and will not include conjecture.

Reporting is next...
Probably the most important part of the whole process is Reporting. It is well known that technical people are usually the last ones to want to document anything. However, in this case, I recommend that you have a template for a forensic analysis report and that you follow it rigorously to ensure that nothing is left out. If a section of the template is irrelevant, simply delete that page from the final report and move on. I find when working with HR and Legal, the analyst’s reporting detail will be greatly appreciated. The more you explain and the more clearly you explain it, the more likely the appropriate course of action will be clear. Do you fire the guy? Or do you discipline him and let him stay as an evangelist to tell others what not to do? Do you have enough information to sue? Or is it better to drop the matter? What is discovered forensically will be important to the outcome of the matter.

So we have taken a look at the Forensics process. How does that relate to eDiscovery?...
Let’s take the basic process of Forensic Analysis and add the nuances and features of the Discovery process. First, let’s take a look at how the eDiscovery process is similar to, and how it differs from, the Forensic Process. First note that all the phases of our NIST Forensic Process are here. Collection, Examination (Processing), Analysis, and Reporting (Presentation). There are, however, other pieces to the puzzle.

Note this chart came from an awesome resource at edrm.net where you can find a well thought-out and reasoned approach to electronic discovery. Let’s take what we have learned about Forensics and tie that in with the EDRM approach to see how that affects our organization.

Notice, by the way, that this diagram does a much better job with the recycling arrows I discussed with you earlier in the forensic process. We can see here that anything discovered at any phase can take the analyst to another parallel phase, or to an earlier phase requiring more Identification or Collection, for example, or what is discovered may push us forward to the next phase. It would be a lovely world if everything were linear, but it simply is not so for most matters.

By the way, if you are looking for a way to evaluate where you are in your eDiscovery program, EDRM.net has a tool...
This eDiscovery Maturity self assessment spreadsheet, just released by EDRM.net, allows you to assign values to the maturity of your program.
You decide between five levels for each of the 7 sections of the self assessment spreadsheet. The 5 levels are:
1. No Process, Reactive
2. Fragmented Process
3. Standardized Process, Not Enforced
4. Standardized Process, Enforced
5. Actively Managed Process, Proactive

Now, back to the eDiscovery process...
The first step of the process is preparation. As with Forensics, eDiscovery requires an Information Governance foundation on which everything else relies. Policies and procedures must be in place. A key policy for eDiscovery is your Data Retention policy.

A good Data Retention policy requires that you know where your data is, and that you have it classified appropriately. Let’s take a quick look at Information Governance.
We’ve already addressed the need for a Computer Forensics Policy. An important piece of the eDiscovery model is the Data Retention Policy. Data Retention can be a nightmare, and for an example I encourage you to spend some browsing time on New York State’s Data Retention Policy site. Here is why Data Retention is a nightmare AND why it is so important for eDiscovery. First, most people really do not know where their data lives. Could you list what systems you have in your organization that have confidential information on them? Can you differentiate between them and the systems with Company Confidential or Internal Use Only information? How long should you keep such data after you no longer use it or need it?

I recently heard a statistic that organizations typically never touch 85% of their information a second time once they have stored it somewhere. That could lead to a whole other discussion about getting rid of data we don’t need, but that isn’t why we are here right now. Most of us are pack rats and we store everything we have forever! Unless you set up a data retention policy. Actually, now that I think on it, a Data Disposal Schedule Policy might be a more appropriate name, since we are really more interested in when we can get rid of it. From a legal Discovery standpoint, every piece of data you have is discoverable. So keeping data just for the sake of keeping it may put your organization at great risk. On the other hand, if you have a defensible data retention policy that says you get rid of everything after so many years, then you follow that policy carefully in practice, then you can remove some risk to your organization by getting rid of that old data.
This sounds simple, but consider how your organization handles email. If you use Outlook, do you also allow PST email archive files? If so, who has them and how old is the mail they store in them? So an email that was sent in 2004 could still live on someone’s computer even though it was long ago eliminated from the main email system or the company email archives, if such exist. Data retention is a huge problem. And it is an integral part of the EDRM Information Governance process.
While Identification doesn’t exist by name in the Forensic process, it is still there. Once there is an issue established for which forensic analysis is needed, the next step is to identify what data needs to be collected from where in order to perform analysis. Is it limited to a single workstation or server? Or are multiple users involved? With eDiscovery, the only difference is that we are locating sources with ESI pertinent to a litigious matter while our Forensics Processes could be wrapped around an HR matter or a Security issue that may or may not turn into a litigious matter. But because there is always that possibility, the collection should be performed in a similar fashion as collections for the Court – that is...
We must gather our ESI in a “Forensically Sound Manner.” That is indeed what will ensure that our ESI is protected as stated here.
Processing is reducing the volume of ESI – that is where, when you are listening to an eDiscovery software company’s sales pitch, you always see the BIG FUNNEL. A WORLD of data goes in at the top, and only what is really relevant falls out at the bottom. That is the whole idea behind reducing volume, whether that is a single matter with a single hard drive that is being examined, or if it is a huge issue involving multiple workstations and servers and email and Sharepoint repositories and databases, and anything else you can think of.

Next the EDRM states we will “Review” for relevance and privilege. Now we are doing more searches, eliminating items that meet certain criteria, and narrowing the responsive data some more.

Then we do analysis which is a bit more detailed in order to find truly relevant data.

Once the data has been culled down as much as is practical, it is time to look at getting that data to the appropriate persons...
Production is more of a technician’s step – where we determine what we have found that could be relevant, then the Forensic Analyst produces that data in the requested form. You may have to produce the data in native, or original if you will, format such as documents and spreadsheets and email files or email archives. You may also be requested to provide the data in a PDF format, or some forensically sound format such as the Expert Witness or Logical Evidence format from EnCase, FTK, or other forensic tools.

Depending upon the size of the data, you may be able to send it to the intended recipient attached to an email, or you may burn it to disk(s) or copy to an external hard drive. In any case, take some steps to ensure that the data is encrypted before you send it by any means to any recipient. Then send the password or passphrase to the recipient Out of Band (by an alternate route), so that it will not be easy for someone to obtain the encrypted data, then decrypt it, unless they are the intended recipient.

The choice of the production format – that is what file type such as native, PDF, or some forensic export type – is negotiated by your Legal team with opposing counsel. Usually opposing counsel will get it in the format they request, unless it can be shown that creating that file type would be an unreasonable burden.

Then the final step of the EDRM process: Presentation...
Presentation is the job of counsel, and technical or expert witnesses, if appropriate. Technicians performing collections, forensic analysts, and other internal Subject Matter Experts (SMEs) can be deposed or called to court as witnesses to explain how they collected data and what they did with the data to prepare it for presentation to the court.

So, how does your organization request or respond in regard to eDiscovery in litigation?...
Even before formal requests for data, according to Lexis Nexis, parties have a duty to Investigate and Disclose types and locations of documents that might be relevant – Data Sources include but are not limited to: workstations, file servers, systems of record, smart devices, network logs, backups, any storage media, etc.

In other words, anywhere data might live, that data is discoverable. (Remember our discussion of the importance of the Data Retention Policy!!)

That being said, there is generally a test of reasonableness parties can expect to be upheld by the court. If obtaining certain data puts an undue burden and expense on one party, they may not be required to produce that data, or, alternatively, the other party may be required to help pay for the acquisition of that data. Take for example emails tape backups that require legacy email systems be reconstructed and backup tapes restored one day at a time in order to find all the stored email between two dates. This can run into $100s of thousands of dollars and resource time from an already time constrained messaging team. Depending upon the information sought, this effort may well not be reasonable.

A paper at the American Bar web site discusses the Rule 26(f) conference where search parameters are negotiated...
In the paper, by Sysol & Huff, called “Slaying the E-Discovery Beast Through an Effective Rule 26(f) Meet and Confer,” there is an extensive discussion around the effort to mutually determine the ESI to be disclosed by both parties in regard to litigation.

The 26(f) “Meet and Confer” has become the eDiscovery opportunity for both parties. Here is where one party may request certain documents of the second party – documents that meet certain search parameters. This is also where Legal will hopefully negotiate with opposing counsel to determine reasonable search terms. The more specific, lengthy, and detailed the search term, the more likely that a relevant document may actually be responsive. Hopefully you will be able to avoid searches for 3 and 4 character words which, because the term is too vague, can return many irrelevant “hits,” and other broad search parameters.

Let’s look at some of the different types of data you may be discovering and the various media upon which it may reside...
Any device that can store data is a candidate for discovery. Everything from micro SD memory cards to huge data servers could have something that is relevant to the matter at hand. Is the case about how a customer was treated? Do you have logs and records stored in SQL or Oracle that may be relevant? Will you need records from an accounting system? What kind of email archives do you have where data may reside? Each of these systems has its own unique challenges for data acquisition, especially when you desire the relevant metadata and that it all be acquired in a forensically sound manner.

Email – probably the greatest source of interest to Legal – can be simple to address if the organization uses a defensible method of archiving ALL email...
For example, there are archive systems available from a number of vendors, such as Symantec, HP, Dell, Barracuda, Proofpoint, and others. You will want an archive system that saves every email message sent or received by your domain, and will store it until the end of the data retention policy period for email, or the end of the litigation hold on the email, whichever is last. Without an archiving system, email will be expensive in time and resources to recover.

Points to consider in choosing an archiving system: Leveraged Cloud vs Internal Hosted, Complex Search Capabilities, additional tools required, ingestion speed, analysis capabilities, export/production options.

For acquiring data forensically, it is best to use a tool designed to preserve all data and metadata without alteration. Let’s look at a few tools…
Here are 4 well known and widely used options for Forensic Software. SIFT was developed under the auspices of the SANS organization and is used in the SANS 408 and 508 courses. Sleuthkit and Autopsy have been around for years and, in fact, were used by SANS prior to the development of SIFT. Both are very capable open source packages. They may, however, present challenges for any organization that encrypts their hard drives, as most do nowadays.

The two commercial packages listed here have built in decryption capabilities. For example, if you use Checkpoint Full Disk Encryption at your organization, EnCase has the ability to decrypt those disks simply by providing the user’s recovery file and the company’s administrative password for Checkpoint FDE. Both FTK and EnCase have the ability to decrypt hard drives encrypted with several different utilities.

Your choice of Forensic software will be based on a number of factors including encryption...
Here is a listing of the current decryption options for EnCase from their web site.

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Product</th>
<th>Supported Versions</th>
<th>64-bit Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Check Point</td>
<td>Check Point Full Disk Encryption (formerly Palisade PC)</td>
<td>6.3.1 up to 7.4, 8.0 (for Windows and Macintosh computers)</td>
<td>Yes</td>
</tr>
<tr>
<td>Crucial</td>
<td>Mobile Guardian</td>
<td>5.2.1, 5.3, 5.4.1, 6.1 through 6.8, 7.3</td>
<td>Yes</td>
</tr>
<tr>
<td>SafeNet</td>
<td>Data Protection (formerly Mobile Guardian)</td>
<td>9.3</td>
<td>Yes</td>
</tr>
<tr>
<td>GuardianEdge</td>
<td>Encryption Plus/Anywhere</td>
<td>7 and 8</td>
<td>No</td>
</tr>
<tr>
<td>GuardianEdge</td>
<td>Hard Disk Encryption</td>
<td>9.1.5, 9.2.2, 9.3.0, 9.4.0, 9.5.0, 9.5.1</td>
<td>Yes</td>
</tr>
<tr>
<td>McAfee</td>
<td>EndPoint Encryption (formerly SafeNet)</td>
<td>4, 5, 6, 7 (for Windows and Macintosh computers)</td>
<td>Yes</td>
</tr>
<tr>
<td>Microsoft</td>
<td>BitLocker and BitLocker To Go</td>
<td>Vista, 7, 8, 8.1, Server 2008, Server 2012 R2</td>
<td>Yes</td>
</tr>
<tr>
<td>Sophos</td>
<td>SafeGuard Easy and Enterprise (formerly BitKiss)</td>
<td>4.5, 5.5, 5.6, 6.0</td>
<td>Yes</td>
</tr>
<tr>
<td>Symantec</td>
<td>PGP Whole Disk Encryption</td>
<td>9.0, 9.9, 10, 10.1, 10.2</td>
<td>Yes</td>
</tr>
<tr>
<td>Symantec</td>
<td>Endpoint Encryption</td>
<td>7.0.2, 7.0.3, 7.0.4, 7.0.5, 7.0.6, 7.0.7, 7.0.8, 8.0, 8.2</td>
<td>Yes</td>
</tr>
<tr>
<td>WinMagic</td>
<td>SecureDec Full Disk Encryption</td>
<td>4.5, 4.6, 5.x and 6.x (including self-encrypting drives)</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Here are the decryption capabilities of FTK.
Nowadays, many software packages have the ability to encrypt saved data files. Microsoft Office will encrypt data, as well as Winzip, 7Zip, and other applications. That being said, there is a difference between “protecting” with a password vs. encrypting. Legacy versions of Microsoft and Zip could “protect” with a password, but the password could be bypassed by forensic tools. However, if the file is truly encrypted with AES or equivalent, the best solution is knowing the password. Otherwise you are looking at “Brute Forcing” the file using some tool. Passware will work with EnCase to decrypt files. Elcomsoft also has tools that will crack passwords on Office files and zip files. Using rainbow tables can help speed up a brute force attack, but if you have a file with an adequate password – like 14 plus complex characters – there is a good chance you will never get into the file.
Hardware Forensic Tool Options – Write Blocking

- **Fred**
- **Dossier**
- **Tableau**
  - [https://www.guidancesoftware.com/products/Pages/tableau/products/forensic-bridges.aspx](https://www.guidancesoftware.com/products/Pages/tableau/products/forensic-bridges.aspx)
- **Software**
  - [http://sourceforge.net/projects/usbwriteblockerforwindows8/](http://sourceforge.net/projects/usbwriteblockerforwindows8/)

Let's look at a few hardware options for forensic tools. Here are links to several well known options. EnCase, by the way, has a write blocking capability built in that is software controlled. Also note the package available at Sourceforge that is a USB write blocker.

Keep in mind the most foolproof way to write block disks is to use a hardware tool such as the Dossier or the Tableau...
The Logicube products and the Tableau products are capable of making write blocked copies of media, as are other available products. The Dossier is capable of cloning or getting a bit for bit image of two drives at once. The tableaus are also capable of imaging a hard drive to a second hard disk. But you can also find equipment that will simply “write block” the media you are imaging and then connect that media to the device making the copy...
Image Masster, spelled with 2 SSs, has a write blocker, as well as other companies out there. This is actually an older unit that has the old style power plugs and works with either IDE or SATA drives. The point is that it is inserted between the computer or device where the image is being collected, and the hard drive being imaged, so any attempt to write back to the source drive will fail, thus protecting the original state of the data.

Side note: With the advent of SSD drives, we have new issues for hashing. Because of defragging algorithms that constantly run on many SSD drives, it is possible the analyst may work on a hard drive that cannot be consistently hashed with MD5 or SHA1. That is because the hard drive is constantly morphing and changing by shifting bytes around from location to location to make the hard drive respond more efficiently. This does not negate the validity of the data, but may require an explanation to the court as to why the data appears to change when actually, it is just being moved around.
Several options exist for forensic workstations. I have seen luggable laptops or suitcase computers as well as towers like the FRED. You can see there are options where various hard drives can be inserted in trays and imaged using the FRED’s built in capabilities. In addition, you may wish to acquire your own equipment and customize it for your purposes. I recommend you purchase equipment preferred by your forensic analyst since they are the folks that will have to use it, and they should have an idea of what they need.

What about Forensic eDiscovery Software? ...
Now, moving from the world of forensics hardware and software to eDiscovery software, I have a few options here for your organization. Consulting is the most expensive option, but it may also be the most expedient method if you do not yet have an internal forensics team. If you have to use a consultant, you may wish to look into having them set up software and tools that they would leave behind for you to use in the future.

**eDiscovery Software**

- EnCase eDiscovery
- Access Data AD Enterprise
- Clearwell
- Some Consulting Firms
  - Mandiant
  - Guidance Software
  - Access Data
  - Kroll OnTrack
  - Symantec
  - Evolver

Now, moving from the world of forensics hardware and software to eDiscovery software, I have a few options here for your organization. Consulting is the most expensive option, but it may also be the most expedient method if you do not yet have an internal forensics team. If you have to use a consultant, you may wish to look into having them set up software and tools that they would leave behind for you to use in the future.
A few notes on network collections using an “eDiscovery” tool vs. imaging entire hard drives. First, imaging the entire HDD is the historical method. However, as hard drives become ridiculously large – say 1TB or more – it is becoming less practical to acquire the entire hard drive image. It is one thing to acquire 250 20 to 40 GB hard drives regarding a particular matter as opposed to collecting the same number of 250GB to 1TB hard drives. It isn’t practical to collect the data, and it isn’t practical to search that amount of space. After all, what the court wants to see is the relevant data and metadata.

If the case is about data on a single hard drive, that is one thing. But if your organization has been asked to provide data from multiple custodians, collecting all the data from all of their hard drives is overkill.

The accepted norm now is to use a forensically sound tool to reach out across the network and collect the data of interest directly from workstations and servers as a kind of snapshot of the data at a point in time, then to store that data in a forensically sound manner so that it doesn’t change once collected. That way we are collecting what we might need and working with that, as opposed to collecting everything in existence and having to cull out the part we don’t need.

That being said, it will still be necessary to do dead drive investigations as well, of course, such as in the case of a terminated employee whose hard drive is no longer in

<table>
<thead>
<tr>
<th>Network Collections</th>
<th>HDD Imaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less impact to production</td>
<td>Large Format Drives a problem</td>
</tr>
<tr>
<td>Users on the network</td>
<td>Better for Criminal Matters</td>
</tr>
<tr>
<td>Depends on bandwidth</td>
<td>Easier to explain and justify</td>
</tr>
<tr>
<td>More Practical (HDD Size!)</td>
<td>Easier collection to ship or pick up drive</td>
</tr>
<tr>
<td>Higher investment up front</td>
<td>Need Secure Storage</td>
</tr>
<tr>
<td>Requires Training / Experience</td>
<td>More time consuming for prep</td>
</tr>
<tr>
<td>Can feed review platform</td>
<td></td>
</tr>
<tr>
<td>Snapshot in nature</td>
<td></td>
</tr>
</tbody>
</table>
use.

Once evidence has been collected, where do you store it?...
Some acceptable options for securely storing collected evidence include any location that can be secured with a key, badge swipe, or other appropriate lock. Depending upon the size of your operation and the number of drives you may collect, you might want to set aside a room with limited access inside of which you might have a locked cabinet or safe. Just remember you need to be able to say unequivocally that no unauthorized person had access to the evidence.

You may even want to institute a double authorization so that two persons are required to log their participation each time some evidence item is accessed.

Physical evidence can also be stored off site at a secure storage location when you have working copies and want to protect the originals.

Lets look at a few options for physical storage...

### Options for secure storage of “Evidence”

- **Physical Evidence**
  - Facility Space
    - Limited Access
      - Keyed
      - Badge Swipe
      - Biometric
    - Safe / Locked Cabinet
      - Keyed
      - Combination
    - Sealed
      - Anti-static
      - Sealed Bags

- **Off Site Storage**
  - Locked
  - Sealed
  - Environmentally sound
  - Controlled Access
  - Auditing
Here is an example of a gun safe that has been modified for storing boxes of hard drives, disks, or other media evidence. This particular safe requires a key to unlock the combination wheel as well as the combination, so it would be simple to institute a requirement for dual authorization to access the evidence if that would be appropriate in your organization.
Here are examples of unpadded and padded anti-static bags. All electronic media, particularly hard drives, should be stored in static protection.
A couple of handy tools of the trade are breakaway evidence tape and evidence bags that seal in such a way that they have to be torn or cut open in order to access the item inside the bag. These are available at a number of sites on line just by searching for the term “Evidence Bags” or “Evidence Tape.” Note that both the tape and the bags have a place for writing CoC information on them for future reference. Any time a seal is broken, there needs to be an explanation of when, why, and who broke the seal. Then once the task is complete – say imaging a hard drive – return the disk to a new sealed bag in which you seal the original bag, so the complete CoC is stored together.
Another option you may wish to consider for storage of evidence files and images is some kind of NAS. I recently toured the T.E.A. in Jacksonville, an energy trading company, and learned that they are using the Dell Compellent system of tiered storage. There are a number of competing systems, of course, from various manufacturers, including IBM and Hitachi, among others.

The T.E.A. indicated that, according to researchers, approximately 85% of all data created is never touched again once it is created. Thus the theory behind a tiered system is to have some Fiber attached storage that is high speed, some medium speed, and iSCSI slow speed storage. Under this system, the high speed expensive storage is used for data that is accessed repeatedly, and those data that are not really being accessed get pushed down to the medium or slow, less expensive, tiers. However for most smaller organizations this will not be practical and the slower speed storage should be adequate. But whether or not you use tiered storage, I recommend you set up replication of the data to a remote site for disaster recovery, depending upon your typical litigation and prediction of losses when your systems fail and you are not able to provide the requested data. (How much is your organization prepared to lose?)
A number of miscellaneous tools are also helpful for forensic work. These can vary depending upon the type of equipment your organization uses and other requirements you may have. Adapters may be necessary for mini or micro SATA, SAS, SCSI II, or special proprietary connections. A Switch may provide a reasonable way to image a drive using network connectivity. Drive docks are essential for easily mounting and dismounting drives for imaging, copying, archiving, etc. A UPS may be crucial to ensure that you do not lose data or damage a disk in the middle of analysis. Having a printer in a secure area may be necessary to keep those who do not have a need to know from observing your output. And one of my favorite tools is the bar code scanner, that allows me to scan serial numbers and model numbers directly off of collected hard drives instead of typing (or mistyping) them.

A few supplies you will need...
1, 2, and 3 TB hard drives are necessary for imaging, and archiving files if you don’t already have an online archiving system. (You can image with a smaller drive as long as it is as large as or larger than the original drive.)

Having USB 3.0, eSATA and other high-speed ports on your workstations is important! With USB 2.0, the best you will ever do is about 1GB per minute when imaging or copying data. USB 3.0 is more than 10 times faster. eSata is as fast as an internal drive – more or less the same as USB 3.0.

I would also recommend, especially if you obtain a bar code scanner, that your organization use the small ½ x 1 ¾ labels for printing your own custom bar codes to stick on devices that you store securely. Bar code fonts can be downloaded for free or purchased. This can make tracking and organization much easier, efficient, and accurate.
A summary of a few documents that are integral to forensic analysis. As previously noted, policy must support the work. Chain of Custody needs to be documented, including a log of evidence and a record of the disposition of evidence. Common procedures should be documented for consistency and defensibility. Investigator’s notes are crucial in the courtroom – the analyst should be allowed access to his or her notes for reference. Forensic Analysis Reports, which should never contain conjecture, are the most important thing the investigator produces.

Side note: that bound composition books with a specific number of pages are perfect for forensic notes and logs because it will be easy to determine if a page has been removed or if the book has been altered in any way.

All this documentation will help provide Legal with what they need...
Legal has “Great Expectations” – especially from the technicians. Collaborate closely with the lawyers to make sure they get what they need and that you manage their expectations. Forensic analysis can take a long time, depending upon what you are looking for. Lawyers should understand that “It is what it is...” and be willing to work with you on focusing the search parameters to reduce your responsive data set and, therefore, their reviewing resource requirements. Review can cost a considerable amount of money.

What Legal wants is everything that might be relevant. What we have to do as analysts is to figure out how to get them just that. What can you do to cull out the irrelevant before delivering data to Legal? How can you suggest refining search terms that might narrow results?

Of course helping them will go better if you understand the case. Ask them questions. Read the complaints and the subpoenas. Find out what they are looking for so that you are looking for the same thing.
So for a quick review... [read the slides...]

- Forensics is a subset of eDiscovery
- Organizations must be ready for eDiscovery requests
- Policy and Written Approval come first, then Process
- Adequate Chain of Custody Documentation, Forensic Soundness, and Preserved Metadata are CRUCIAL
- NIST SP800-86 = Forensic Processes
- Edrm.net = eDiscovery Processes
- Counsel could request “the moon” in 26(f) eDiscovery “Meet and Confer”
Review (cont.)

- All Data is Discoverable, although there can be a reasonableness test, and privileged communications
- Forensic Software is available as open source or commercial
- Endpoint Encryption may dictate your forensic tool
- Collection may be performed locally or across the WAN
- Collected Physical and Logical Evidence must be stored securely
- Your mission is to keep the lawyers happy!
Review – the last word

- Everything must be done in a forensically sound manner
- Preserve all evidence with Chain of Custody
- Document Document Document
Thanks for the opportunity to meet with you today. Please feel free to contact me via email. I hope this has been helpful information for you.

If you have any questions, now is the time to ask them.