

**The Dangers of Do-It-Yourself eDiscovery**

**July 2008**



# Table of Contents



Introduction **page 3**

Understanding the Dangers of Do-it-Yourself **page 4**

A Safer Approach to eDiscovery: Outsourcing **page 7**

How to Select an Outsourcing Partner **page 8**

Conclusion **page 10**

# The Dangers of Do-It-Yourself eDiscovery

*Electronic Discovery (or eDiscovery) is a hot topic in the boardrooms of most organizations today, and for good reason. The risk associated with failing to preserve electronically stored information (ESI) – not only from existing active network shares and email systems, but also from backup media residing in storage – is severe. Today’s news is filled with stories of fines, sanctions, or even settlements of frivolous lawsuits simply because an enterprise was unable to find and produce data relevant to a given case. Thus, the question in front of most executives isn’t if they will require eDiscovery products and services, but rather, when litigation arises, what method of response will the organization choose?*

## Introduction

For most enterprises, the aspect of eDiscovery that poses the greatest risk is collecting, preserving, and processing historical data that is sitting on backup media. This information, which may equate to terabytes of data on tens of thousands of tapes, is regularly requested in court.

and liabilities associated with the handling of electronic evidence. A partial DIY solution may also be employed by outsourcing a portion of the eDiscovery requirements while taking ownership of the remainder of the project in-house. Typical DIY solutions may include out-of-box appliances, add-on software applications, and independent data extraction technologies.



*According to our research, a majority of corporate counsels say that data requested as part of an electronic discovery inquiry is 36 months or older.*

*Brian Babineau  
Enterprise Strategy Group*

In order to get a better handle on this risk, most organizations are taking one of two approaches:

- **In-House “Do-It-Yourself” Method** – Some organizations opt to purchase, staff, and implement an internal do-it-yourself (DIY) offering. In this case, the enterprise bears all processing responsibilities

- **Outsourcing Method** – Rather than assume the responsibility internally, an organization may partner with a service provider that is proficient in all phases of eDiscovery. Typical services provided by this type of vendor include litigation planning, data preservation & collection, data processing, analytics & review, and production. The outsourcing method is generally the preferred approach by enterprises that don’t have the internal expertise and processes to conduct eDiscovery – and ensure data defensibility – on their own.

Traditionally, the vast majority of enterprises have favored the outsourcing method, in large part because these organizations lacked the in-house capabilities to conduct eDiscovery efforts on their own. However, as technology has improved and organizations have gotten savvier about eDiscovery, some executives are starting to consider the DIY approach due to perceived process efficiencies and cost containment potential.

# The Dangers of Do-It-Yourself eDiscovery

Beware. The DIY approach can be cost-effective for small projects that only involve basic backup media, but this method also contains many hidden dangers. This white paper will outline those hidden dangers to help you make a more informed decision on how to address your eDiscovery challenges.

## *Understanding the Dangers of Do-it-Yourself*

To help you better understand the risks associated with DIY eDiscovery, we will break down this approach by looking at three basic functional areas: people, process, and technology. In other words: you want to make sure you have the right people using the right tools in the right way.

### **People**

A critical component of eDiscovery is the personnel assigned, as success is dependent on their expertise and experience. Poor personnel assignments will cause efficiency and efficacy challenges. As you evaluate a DIY approach, make sure you pay particular attention to:

### **Industry Expertise**

Quite often, at the heart of any DIY project is the assumption that a technological solution will “do most of the work” and the internal staff will simply execute and monitor the process. This fallacy unfortunately dismisses the level of expertise required of those actually handling the media and engaging the technology.

The reality is, those working on eDiscovery projects must understand the legal requirements and ramifications in play, the hardware being engaged for data handling, the types and structure of the data involved, the intricacies of the software being employed, and most importantly, respect every step in the eDiscovery process to ensure each phase is conducted appropriately.

### **Scalability**

When litigation arises and deadlines are tightened on a moment's notice, the ability to identify and engage internal personnel with ample expertise -- while simultaneously pulling all other necessary resources together -- is exceedingly challenging. When organizations opt for DIY methods, locating qualified individuals with requisite experience in pertinent eDiscovery matters can be problematic. Many times, corporations rely heavily on already thinly stretched IT personnel to locate and process ESI for litigation purposes.

### **Process**

A fundamental series of systematic actions designed to provide consistency and defensibility for handling, tracking, and reporting on electronic evidence throughout the lifecycle of the project are key components of eDiscovery. Specific process-related dangers associated with the DIY method include:

### **Chain-of-Custody**

Chain-of-Custody (CoC) is defined as “the documentation and testimony regarding the possession, movement, handling and location of evidence from the time it is obtained to the time it is presented in court.” It ensures there is an unbroken trail of accountability for all media and the ESI derived from it throughout the processing lifecycle.

In order to track and prove CoC, an organization must have rigorous and consistent procedures in place, including documentation for all phases. This means keeping precise records of all pertinent information (media type, labels, serial number, origin, and the location where it is to be stored, etc.). Specific questions you'll need to answer include:

- Where is the media and who has it?
- When was the media checked back in?
- What data was extracted from which media?
- Can a clear line be drawn from file to source media to technician?

# The Dangers of Do-It-Yourself eDiscovery

Often, organizations that elect to conduct DIY eDiscovery find they do not have the processes in place to track the above with the precision that is required. As a result, CoC can be questioned and all the evidence that your organization has worked so hard to prepare may not be legally defensible – possibly leading to expensive ‘rush-job’ work or even fines and sanctions.

## Media Tracking

Media tracking is a methodology for ensuring that each piece of media has gone through the needed steps for processing. Without sound media tracking, it is likely media can “fall through the cracks” of the defined process, which results in incomplete and indefensible results. The best media tracking methods tie into the CoC. Remember, it is not enough to know a piece of media was checked out to be processed; it is also necessary to know that the media was processed.

In general, organizations that conduct DIY eDiscovery rely on either homegrown or third-party applications to track whether or not ALL media was processed. The problem is that these applications usually do not come ready to integrate into a CoC system to provide instant feedback when a piece of media is actually processed. Thus, it is incumbent on the organization to create a media tracking methodology – usually a manual system that is inefficient, time-consuming, and highly susceptible to human error. (An automated system that is updated in real time is far more efficient and accurate, but it isn’t standard with off-the-shelf technology.) Regardless of your approach, without a media tracking system that is airtight – before processing begins – your organization is at great risk of mistakes, rework, and time loss in the near future.

## Data Tracking

Data tracking begins when the data is inventoried from the media and continues throughout the processing lifecycle. During all phases of production, for defensibility, each file must be reconcilable and data integrity must be maintained. This is where data tracking is paramount and

problematic. Tracking, reconciliation, and processing may be required for billions of pages of output.

Most, if not all, DIY solutions are unable to handle this volume of data, and the potential exists for production issues, lost time, and inaccuracies. As with media tracking, if the pertinent data is never presented to the technology to be processed, the technology is useless.

## Quality Assurance

Quality Assurance (QA) is the process of verifying the accuracy and integrity of ESI throughout the eDiscovery lifecycle. It also includes validating that every piece of media underwent all phases of processing.

### QA utilizes

media tracking to determine where each piece of media was stored, whether it was processed correctly, and that all pertinent data from the media was extracted to completion. Through data tracking, QA reports on every file and reconciles it to its original media, whether the file was culled, processed via data migration, de-duplicated, and/or keyword searched. Audit logs should be readily available and reports should be generated rapidly to provide certification of work and validity of output.

To ensure defensible, high-quality data, QA must be able to verify every step of the process from both a CoC standpoint as well as a data processing standpoint. Again, in the absence of pre-packaged processes, the DIY approach often falls short in QA as it puts a strain on resources that are not accustomed to managing data as “evidence.”

## Technology

Technology comprises the tools and methodologies utilized to perform a specific technical function within the scope of the project. At first, it is the easiest portion of the job to budget. However, there are many additional “hidden costs” and risks that often surface with DIY

# The Dangers of Do-It-Yourself eDiscovery

tools that are not anticipated. Below are some areas where you can look for them:

## Legacy Backup Media & Software

Backup media that has accumulated through merger and acquisition activity or that has been replaced by larger and faster media types typically contain historical data that was written with older versions of potentially obsolete software unsupported by most DIY eDiscovery applications. Even if one media type was consistent across an inventory of legacy data, the likelihood of the same version of the same backup software being used for years is incredibly low. Generally, most of these problems do not crop up until late in the eDiscovery process using the DIY approach, leaving the organization with the choice of either skipping potentially relevant information or paying for expensive “rush job” extraction by a specialist.

## De-duplication

De-duplication, or the removal of forensically identical files and email messages, is a standard eDiscovery procedure intended to reduce overall volume while maintaining relevant data. However, not all de-duplication is the same. DIY applications sometimes sacrifice accuracy for speed, and custom tools sometimes disregard industry standards. At best, the result can be more data to go through the costly process of review. At worst, your data becomes much less defensible. Before launching any DIY project, understand how the de-duplication tool hashes its information, and understand how to defend the results.

## Foreign Characters

Often, foreign characters – data encoded with a form atypical to the base install of the technology – will emerge in a dataset. (This is true even for organizations doing all their business in the U.S.) Many DIY eDiscovery applications are not built to handle this complexity, causing them to miss or misinterpret some relevant information. These common mistakes associated with DIY have been known to lead to great expenses in the future.

## Indexing/Keyword Searching

DIY Indexing and keyword searching are areas that may deliver a substantial amount of variance. The number of options and settings on many third-party indexing tools can be staggering, and the slightest alteration of these settings can have a massive impact on how the ESI is indexed and how the resulting output is generated. Because of the high potential to accidentally alter the data set, this approach lends itself to incredible risks – including misclassification of data as “privileged” or “non-privileged” evidence. Before engaging a DIY application, be aware of the significant room for error – inadvertent or otherwise – during configuration and installation of the tool.

## Antiquated Email Stores

Because organizations are continually upgrading, replacing, or consolidating outdated email systems, the volume of “historical” ESI is growing quickly at many enterprises. This is bad news for DIY applications that often do not support older email formats. In many cases, this data is simply overlooked by the third-party tools that are being used by an untrained eye. Again, organizations should be cautious about using DIY applications for eDiscovery collection and processing if they are serious about capturing every single piece of potentially relevant data.

## Media Extraction

DIY applications developed for media extraction fall into two categories: native and non-native.

Native DIY extractions require delicate re-creation of the original (i.e. native) environment before the data can be extracted from the media. Because this process is typically slow, cumbersome, and riddled with complexity, the amount of overhead required to actually produce the data is sizeable. Further, most DIY applications are not designed or equipped to re-create the original environment, so an in-house approach will almost always fall short in this respect.

# The Dangers of Do-It-Yourself eDiscovery

Non-native DIY extractions do not require the original environment to be re-created before data extraction can begin, but the process can be tricky. There is enormous potential to miss data when performed by personnel with less-specialized knowledge and experience. Relying on DIY tools for non-native extraction can also cause an eDiscovery project to come to a complete halt, should the DIY application require upgrades or new versions in order to support different media types.

## Data Storage

Data storage is a significant concern when dealing with any DIY eDiscovery project, and it often presents considerable costs that are not uncovered until after the fact. Make sure you analyze your storage requirements for the project in advance, along with the best way to optimize storage usage. As with most storage-related concerns, the answer is rarely limited to just the hardware supporting the technology; there are also network, security, and infrastructure complexities and related costs.

Also note that predicting the amount of storage space needed for an eDiscovery project is often a challenge. Organizations that adopt the DIY approach, and purchase (or have available) only a limited amount of storage, are setting themselves up for a big surprise in the future – particularly if multiple eDiscovery requests surface at the same time.

## ***A Safer Approach to eDiscovery: Outsourcing***

At first glance, outsourcing to a third party may appear expensive when compared to DIY eDiscovery offerings. Other potential concerns about this approach might include reservations about the risk associated with transferring tapes to/from vendor facilities, or the possibility of longer turn-around times since these facilities are not on-site.

However, outsourcing is in fact a smarter and safer alternative for most organizations. These trusted partners have the people, processes, and technology necessary to protect their customers from the above dangers throughout all phases of the eDiscovery process – from initial planning, data collection, processing, production, and document review.

## BENEFITS OF

### OUTSOURCING EDISCOVERY

#### People

- Strong legal and technology pedigree
- Project management and status reporting included

#### Process

- Sound processes to track and defend chain-of-custody
- Media tracking, data tracking, and QA built into the offering
- Experts can provide written and oral testimony to defend the processes used

#### Technology

- Comprehensive support for a variety of media and software
- No data skipping due to de-duplication or foreign character deficiencies
- Ability to scale as necessary for growing scopes of work
- Avoidance of spoliation and adverse inference
- Ability to manage multiple matters concurrently

# The Dangers of Do-It-Yourself eDiscovery

Before your organization makes a decision between taking eDiscovery in house or outsourcing, consider the following:

## People

When responding to a discovery order, why risk placing your data in the hands of internal staff who are either inexperienced with eDiscovery tools, uninformed about the liabilities associated with handling ESI as “evidence,” or too burdened by their existing workload to provide the attention that the eDiscovery process demands? Outsourcing to an eDiscovery firm takes the burden of human resources off your plate by providing the project management and eDiscovery expertise you need.

## Process

How comfortable is your organization with assuming full liability for the methods of data collection and processing conducted internally, rather than sharing the risk with an eDiscovery expert? A sound eDiscovery vendor will provide legally defensible processes, in addition to quality assurance measures, to ensure security and accuracy at all times.

## Technology

Is your enterprise willing to accept the possibility of missed data, erroneous processing, or unsupported media/software formats... after the fact? External eDiscovery vendors are designed to scale to meet your needs, while offering comprehensive support along the way to ensure a complete, exhaustive approach to your eDiscovery needs.

Ultimately, by outsourcing eDiscovery instead of DIY, organizations are dramatically increasing the defensibility of their data, eliminating big surprises, and decreasing their total cost of eDiscovery.

## *How to Select an Outsourcing Partner*

Now that you have learned about the pitfalls associated with DIY eDiscovery and heard the benefits of outsourcing these important processes to an expert, you are probably wondering how to select an outsourcing partner to help protect your organization from eDiscovery risk. In general, a sound service provider has the industry expertise, technology, defined processes, physical security, and ability to deliver pertinent ESI in a way that is reconcilable and defensible. They also can deliver a full complement of eDiscovery offerings, including data archiving, collection, processing, review and production.

Using the criteria above, you should be able to narrow down your search to a handful of potential service providers. Your final decision should hinge on what you learn during the advanced due-diligence stage. During this stage, you'll want to learn more about each of the following items:

## History & References

There are a growing number of new eDiscovery providers in the market. Many of these newcomers use off-the-shelf tools for data processing, which can lead to significant (and costly) errors, particularly when operated by inexperienced employees. Therefore, it is critical to understand more about the legal and IT expertise of your potential vendor. Be sure and ask for customer references -- in an industry and with a similar volume of data -- that is comparable to your situation. You may also want to ask how long the vendor has been in business, and for a breakdown of technical staff, project management, and legal expertise to ensure sensitive information is appropriately processed and protected.

---

# The Dangers of Do-It-Yourself eDiscovery

---

## Capabilities

Generally, submitting a Request For Proposal (RFP) to evaluate a potential vendor's capabilities is an appropriate method of determining whether or not that particular vendor will provide a good fit for you. If the RFP process is not feasible (e.g., due to time constraints), preliminary meetings can provide a quick, less-formal method of evaluating a vendor's strengths and weaknesses. In either case, be sure you walk away understanding how much capacity the vendor has (in case the scope of your project expands), the security measures taken at their data center, the processes taken to ensure all data has been extracted, and how they track chain-of-custody from the minute it arrives at their facility. Lastly, if time allows, engaging a potential vendor for a "sample project" can provide considerable insight to help determine with certainty whether a particular vendor understands the project requirements and can deliver exactly what is needed.

## Project Management

A project manager can make or break your project. When seeking out vendors, look for one with a solid project management team that understands your problem and communicates effectively. Proactive communication through daily or weekly status reports is a bare minimum. Also watch out for the bait-and-switch approach, whereby a seasoned project manager is introduced early on but a more-junior project manager takes over after the initial sale. Of course, identifying project managers who have delivered prior testimony, court opinions, or affidavits is always valuable, just in case you are required to defend your eDiscovery methodology in Court.

## Facilities

Make sure you visit your selected vendor's facilities before finalizing your decision. With data that carries as much risk as ESI, you cannot risk trusting your potential evidence with a company that does not have government-level security, technical prowess, and chain-of-custody tracking down to a science. Some of the specific things you should look for include: disaster recovery capabilities, redundant safety measures, quality control methods, 24-hour support, personnel identification measures, entry and exit manifests, updated CoC logs, and restricted access to client data.

---

# Conclusions

---

## *In Summary*

Everyone intellectually understands eDiscovery is not an easy process, yet some organizations are regularly fooled into taking a do-it-yourself approach, thinking that technology has evolved far enough to eliminate the majority of that complexity. Nothing could be further from the truth. The collection and processing of ESI alone requires an optimal mix of people, process, and technology that most organizations can't bring together.

Failure to recognize that eDiscovery is not your organization's core competency and ultimately taking a DIY approach can lead to several dangers. These include:

- Substantial business disruption, loss of time, and a drain on resources
- Data Process, security, defensibility, quality control, and CoC are diminished
- Scalability can be incredibly challenging when tightened timeframes are involved
- Data extraction from outdated or damaged media and obsolete software is often impossible
- Relying on internal IT departments spreads resources even more thinly
- "Native" restoration requires difficult re-creation of original environment
- Internal human error can be catastrophic

A better approach to eDiscovery, particularly when historical data is involved, is to outsource to a trusted third party. When selecting such a service provider, pay careful attention to a firm's capabilities, background, and expertise. Also look for a vendor that has the people, processes, and technology to serve all of your needs and protect you from the above dangers. In the end, you will enjoy a more efficient and cost-effective eDiscovery process, while taking clear steps toward ensuring the defensibility of your data.

---

## About RenewData

---

RenewData is a leading provider of e-discovery and ESI (electronically stored information) risk management services to assist corporations and law firms responding to lawsuits, investigations, and audits. Superior legal expertise, scalable technology, and a state-of-the-art facility featuring government level physical security enable RenewData to provide clients with secure, accessible, and manageable data in a cost effective and timely manner. RenewData's e-discovery services cover the five critical steps of the e-discovery process, including planning, preservation and collection, processing, review, and production. RenewData's ESI risk management services, which include backup tape liability management, data migration, and evidence storage, provide corporations with a proactive means of managing the risks associated with ESI. RenewData has been ranked a "top provider" in the Socha-Gelbmann Electronic Discovery Survey Report for three consecutive years and was ranked in the 2006 and 2007 Inc. 500 lists of fastest-growing privately held companies in the U.S. For more information, visit [www.renewdata.com](http://www.renewdata.com) or call 888.811.3789.



512.276.5500  
888.811.3789

[www.renewdata.com](http://www.renewdata.com)

Copyright © 2008 Renew Data Corp. All rights reserved. RenewData and ActiveVault are registered trademarks and ESIRM and "Single-Pass Processing" are trademarks of Renew Data Corp. RenewData disclaims any proprietary interest in the marks and names of others.