



THE E-DISCOVERY “HOW-TO” GUIDE:

Practical Recommendations for Streamlining Corporate E-Discovery

A Clearwell White Paper

Table of Contents

Introduction	3
The Challenge: The "E-Discovery Squeeze Play".....	3
The Solution: Manage E-Discovery as a Business Process	3
Key Components of E-Discovery.....	4
I. Identification, Preservation, and Collection.....	4
<i>Best Practices</i>	5
<i>Supporting Technologies</i>	5
II. Processing, Analysis, and Review.....	6
<i>Best Practices</i>	6
<i>Supporting Technologies</i>	6
III: Production and Presentation	7
<i>Best Practices</i>	7
<i>Supporting Technologies</i>	8
Where to Begin?	8
Conclusion	9
E-Discovery Resources	10

Introduction

The amended Federal Rules of Civil Procedure (FRCP), recent headlines highlighting spoliation sanctions, and rising costs have made e-discovery a top priority for legal, IT, security, and records management teams. However, enterprises are not sure where and how to begin improving their e-discovery process and often struggle to generate a shortlist of technologies to help them. This paper breaks down the e-discovery process into phases, discusses best practices and relevant technologies for each phase, and provides a matrix to guide an e-discovery team's decision making process.

THE CHALLENGE: THE E-DISCOVERY "SQUEEZE PLAY"

Lawsuits, corporate investigations, and regulatory audits are increasing in number—rapidly. In 2006, the average number of lawsuits at large enterprises reached 556, up 270 percent from the previous year.¹ To further compound matters, the amount of electronically stored information is also dramatically increasing. The Sedona Conference estimated that 93 percent of all the information generated by businesses in 2006 was produced in electronic form.² On top of it all, the recent amendments to the FRCP have brought e-discovery to the forefront of every legal matter, both in Federal and increasingly in State courts. As a result, e-discovery teams consisting of legal, IT, security, and records management professionals are getting squeezed as they struggle to cope with increasing workloads while attempting to control spiraling e-discovery costs. Organizations that cannot master this difficult process are putting the outcomes of matters at risk.

THE SOLUTION: MANAGE E-DISCOVERY AS A BUSINESS PROCESS

E-discovery is a fact of life in today's business environment, and yet many companies still fail to address this critical task in a proactive manner. Just like disaster recovery and security, enterprises should realize that e-discovery is inevitable, and the best option is to prepare for it. However, proactively addressing e-discovery and implementing a set of repeatable processes and technologies to streamline the numerous, interrelated tasks is not a trivial matter. E-discovery is not a single activity, but a series of steps, involving the coordination of multiple disciplines across the enterprise. This complexity makes tackling the entire process at once ill advised. Instead, a more effective approach to improve your e-discovery process is to adopt a “divide and conquer” strategy and focus on the steps that deliver the biggest impact first.³ In this fashion your organization isn't faced with a “boil the ocean” problem of trying to do too much at once. Accordingly, an iterative approach allows the team to start to measure pertinent metrics that will objectively illustrate the impact of improvements.

¹ 2005/2006 Fulbright and Jaworski Survey, available at <http://www.fulbright.com/>

² ARMA, “What is Electronic Records Management (ERM)?”, <http://www.arma.org/rim/101/articles.cfm?key=rim101erm>

³ “Impact” will probably vary per organization and can be thought of as a combination of (i) risk reduction due to the avoidance of spoliation sanctions attendant with e-discovery abuses, (ii) cost reduction, which is comprised of the reduction in e-discovery processing costs and attorney review costs, and (iii) reduction in organizational demands since a complex e-discovery matter may consume cycles legal, IT, security, and records management teams..

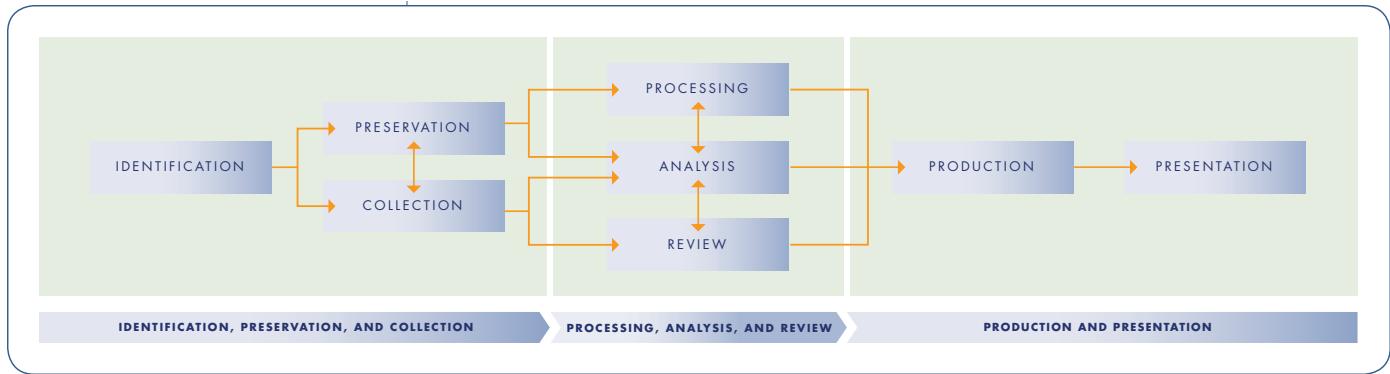


Figure 1:
Simplified Electronic
Discovery Reference Model.
Source: EDRM Project
(<http://www.edrm.net>)

KEY COMPONENTS OF E-DISCOVERY

The best way to initially understand the entire e-discovery process is to use the architecture and glossary promulgated by the Electronic Discovery Reference Model (EDRM) Project. A simplified version of the EDRM process flow is pictured in Figure 1. Launched in May 2005, the Model provides a common, flexible and extensible framework for the development, selection, evaluation, and use of e-discovery products and services.

When responding to an event, such as a litigation or a regulatory inquiry, the e-discovery process consists of three major phases:

1. Identification, Preservation, and Collection
2. Processing, Analysis, and Review
3. Production and Presentation

Each of these major phases has specific requirements and costs. It is estimated that enterprises spent over \$1 billion to identify, preserve, and collect ESI in 2007. Furthermore, enterprises spent almost \$2.5 billion to process, analyze, and review ESI (making it the most expensive phase of e-discovery) and \$576 million to produce it.

I. IDENTIFICATION, PRESERVATION, AND COLLECTION

The identification, preservation, and collection steps of e-discovery are triggered in litigation context once litigation is “reasonably likely”⁴ but can also be brought into play by a judicial order or governmental inquiry. Regardless of the impetus, the first leg of the journey (i.e., the identification step) refers to the detection of all the electronically stored information (ESI) that needs to be preserved and potentially collected. In many instances the preservation obligation is broader than the duty to collect and subsequently produce ESI. In either case, the goal is almost universally to identify (1) the custodian(s) or key player(s), (2) the appropriate date ranges, and (3) the sources of ESI. During preservation the goal is to institute a legal “hold” to ensure that potentially relevant/responsive ESI is not accidentally destroyed or lost as a result of the routine, good faith operation of an electronic information system.⁵

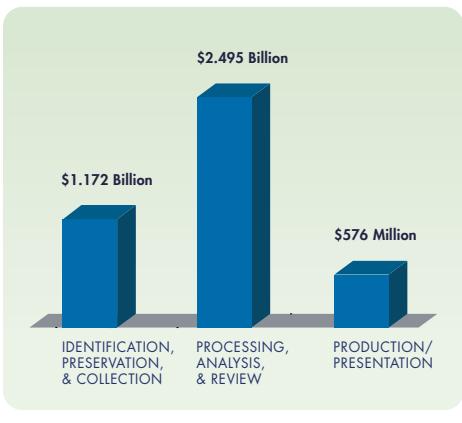


Figure 2:
Cost of Major E-Discovery Phases
Source: IDC Research, Gartner,
Clearwell

⁴ The Sedona Conference Commentary on Legal Holds, Working Group on Electronic Document Retention & Production, 2007

⁵ See, FRCP 37(f) (renumbered as Rule 37(e)) which provides a limited safe harbor if certain precautions are taken.

Finally, the collection step refers to the acquisition of ESI marked as potentially relevant in the identification phase. Together, the identification, collection, and preservation e-discovery steps cost enterprises over \$1 billion in 2007.⁶

Best Practices

The following best practices have been employed by e-discovery teams at leading companies for identification, preservation, and collection of ESI:

- Be as thorough and comprehensive as possible during the identification step. Begin by interviewing individuals who were directly involved in the disputed matter, as they are likely to have the majority of the information relevant to the matter. They are also likely to provide insight into other individuals who may have been involved in the matter.
- Make sure that all individuals identified are informed of the importance of the litigation "hold," how it should be enforced, and who to contact with questions. Be sure to discuss the consequences of not adhering to the litigation hold with each person receiving the litigation hold notice.
- Use forensically approved technologies to collect ESI from different data sources, as no single product collects ESI from all data sources. Implement processes that are aligned with the different tools to ensure comprehensive collection, with necessary indices of authentication and chain of custody. See below for technologies that provide identification, preservation, and collection solutions.
- Document your litigation hold and collection processes. Create and maintain a checklist for every case to make sure that you are following your documented processes. Consider hiring a third-party firm to audit your litigation hold and collection processes to ensure they are sound and defensible.

Supporting Technologies

Based on the Socha-Gelmann 2007 Electronic Discovery Survey results, as well as reports from IDC Research and Gartner, the leading identification, preservation, and collection technologies include:

- Forensics products such as Guidance EnCase[®] and Paraben's Forensic ToolKit[®]. These products are typically used to create forensically sound copies of ESI on files shares, user desktops, and laptops. These solutions provide pre-filtering features to collect ESI by searching for custodian, file types, and date ranges.
- Archiving vendors, such as Symantec Enterprise Vault[®] and HP Integrated Archive Platform[®]. These products excel in identifying and preserving evidence from the data (typically email) stored in their archives. As with forensic tools, these solutions typically provide pre-filtering features to collect ESI by searching for custodian, file types, and date range.

⁶ Combined research from IDC, Gartner, and Clearwell.

⁷ Dataquest Insight: E-Discovery Market Drives New Search, Content and Records Management Investments, Gartner, Nov 2007

II. PROCESSING, ANALYSIS, AND REVIEW

The primary objective of this next phase is to review user data, analyze the case, and identify relevant and privileged ESI. The processing step refers to the crawling, indexing, and culling of ESI, including any conversion of ESI that enables analysis and review to be done electronically. The analysis step refers to the assessment of ESI for relevant summary information, such as key topics, people, and timelines critical to the case. The review step refers to the detailed review of the ESI and evaluation of ESI for relevance and privilege. This step is usually performed in conjunction with internal and outside counsel. As the penultimate phase, these steps are both expensive (in terms of processing and attorney review hours) and critically important because this phase directly impacts the early case assessment and overall case strategy—such as the decision to fight versus settle—and ultimately, the outcome of the case. In addition, privileged documents that are not identified correctly and accidentally turned over may often not be retrievable.

Gartner estimates that the average e-discovery “event” costs \$1.5 million, of which the majority covers expenses accrued during the processing, analysis, and review steps of e-discovery.⁸ In fact, this phase cost enterprises approximately \$2.5 billion in 2007.⁹ The high costs associated with these steps include:

- Processing Fees: Service providers may charge \$2,000 or more per GB to process ESI
- Legal Review: Outside counsel fees range from \$200 to \$300 per hour to manually review ESI, meaning that attorney review fees for a gigabyte of data could be as much as \$30,000.¹⁰

Best Practices

The following best practices have been employed by e-discovery teams at leading companies to improve reliability and lower the costs of processing, analysis, and review of ESI:

- Prioritize data “culling” (i.e., reduce the total size of ESI to be processed or reviewed), as it can have a significant impact to the cost and duration of e-discovery. According to law.com, 90% of the data collected is irrelevant to the case, yet companies continue to send most or all of the data to service providers at the average cost described above. Significant cost savings can be achieved by leveraging technologies to remove duplicates and pruning irrelevant data and false positives. As a result, forward thinking enterprises only send the relevant 10% to service providers for processing.
- Focus on early case assessment strategies to understand scoping information about your case. Information about critical custodians, time line, events, and data types will be extremely valuable to make important early decisions and will significantly improve the productivity of all downstream electronic discovery activities.
- Leverage analytical technologies to your advantage during meet and confer e-discovery sessions. For example, instead of blindly negotiating “keywords,” conduct keyword searches and use actual search results to convince opposing counsel to agree with your recommendations. Additionally, use data sampling techniques to define cost and burden assumptions.

⁸ Meeting Compliance and E-Discovery Demands: What IT Managers Need to Know, Gartner, 2006

⁹ Combined research from IDC, Gartner, and Clearwell.

¹⁰<http://www.abanet.org/lpm/ltt/articles/vol2/is1/ediscovery-early-case-assessment.shtml>

- Use contextual review to reduce review costs. The traditional approach to reviewing—keyword hit by keyword hit, one document at a time—may result in redundant effort, greater costs, and inconsistent tagging. Next generation e-discovery software products provide advanced analysis capabilities such as topics and discussion threads to show evidence in context and dramatically accelerate review. For example, most email messages are part of a longer discussion thread, and it is much more efficient to review the entire thread at once, versus one message at a time in random order.

Supporting Technologies

Based on the Socha-Gelmann 2007 Electronic Discovery Survey results, in addition to reports from IDC Research and Gartner Group, the leading processing, analysis, and review technologies include:

- Next-generation e-discovery review and analysis products such as those from Clearwell Systems and Recommind.¹¹ These solutions often offer new and advanced capabilities such as discussion threads, relevance rank, topics, auto-filters, nested searches, one-click tagging, and streamlined review workflow. Many of these newer technologies also export data in standard formats, such as EDRM XML, for easy import into traditional review platforms.
- Search vendors such as dtSearch®.¹² These vendors provide advanced search capabilities, but they typically lack any analysis features and do not provide legal or case workflow capabilities.
- Traditional review platforms such as Concordance and Summation.¹³ These products provide the ability to review one document at a time.

III: PRODUCTION AND PRESENTATION

The final phase of the e-discovery process, production and presentation, focuses on preparing and producing the ESI for the courts or regulators. The production step refers to delivering ESI to various requesting/propounding parties in various media and formats. The presentation step refers to displaying the electronic data at depositions, hearings, trials, etc., in order to elicit further information or persuade the audience. Together, production and presentation cost enterprises approximately \$576 million in 2007.¹⁴

Best Practices

The following best practices have been employed by e-discovery teams at leading companies for production and presentation:

- Ensure that production details, such as form, media, and schedule are negotiated and agreed upon by both parties. Clearly communicate your capabilities and limitations, data formats, and any other technical details to your counsel so they can be addressed during negotiations. This will eliminate unnecessary re-work and late-night fire drills.

^{11, 12, 13} Dataquest Insight: E-Discovery Market Drives New Search, Content and Records Management Investments, Gartner, Nov 2007

¹⁴ Combined research from IDC, Gartner, and Clearwell.

- Gain clear commitment on who will perform the work required during production—whether it is a third-party service provider, outside counsel's litigation support department, or in-house litigation. Be sure to understand the production requirements and map them to the capabilities of the various parties. For example, an in-house litigation support department is likely to produce the data in native file format along with a privilege log, while a service provider is best suited to handle the complex production requirements that involve redaction and bates stamping.
- Track the history of document productions—both paper and electronic-based—and make sure all are documented by each person or company that handles media as it moves through the processing and discovery process.
- Always evaluate the production documents received, as well as those being produced. Immediately upon receipt of production media, conduct a preliminary review of the data and provide a memo to the lead attorney, paralegal, or team about what was received and an assessment as to completeness and ability to load the data.
- Continually review and sample productions to ensure that privileged or confidential files are not inadvertently produced to the opposition.

Supporting Technologies

Based on the Socha-Gelmann 2007 Electronic Discovery Survey results, as well as reports from IDC Research and Gartner Group, the leading companies in this space include CT Summation and Dataflight Concordance. These products provide similar capabilities, so most organizations simply choose the product used by their e-discovery service provider or law firm.

WHERE TO BEGIN?

As companies streamline their e-discovery processes to react more quickly, control costs, and ensure compliance with the amended FRCP guidelines, it is to their advantage to develop repeatable processes that meet their needs in the most cost-effective way, while mitigating potential risks.

How do you do this? Best practices in e-discovery dictate that e-discovery teams follow several actions (also depicted in matrix form in figure 3):

- Audit and document your legal hold and collection processes. The need to collect and preserve ESI once litigation is reasonably likely is fraught with danger. If you have dealt with litigation before, it is likely that you have some processes in place. If these processes need improvement, weigh it against the opportunity cost of streamlining the more expensive phase of processing, review, and analysis.
- Perform early case analysis, where appropriate. Can you easily discover the top custodians and date ranges for specific keywords? Can you identify additional custodians that you might have missed? Can you find the “smoking gun” that will significantly influence your case strategy? Can you quickly perform keyword searches to guide keyword negotiations to your advantage during meet and confer sessions? You should be able to answer “yes” to all of the above.

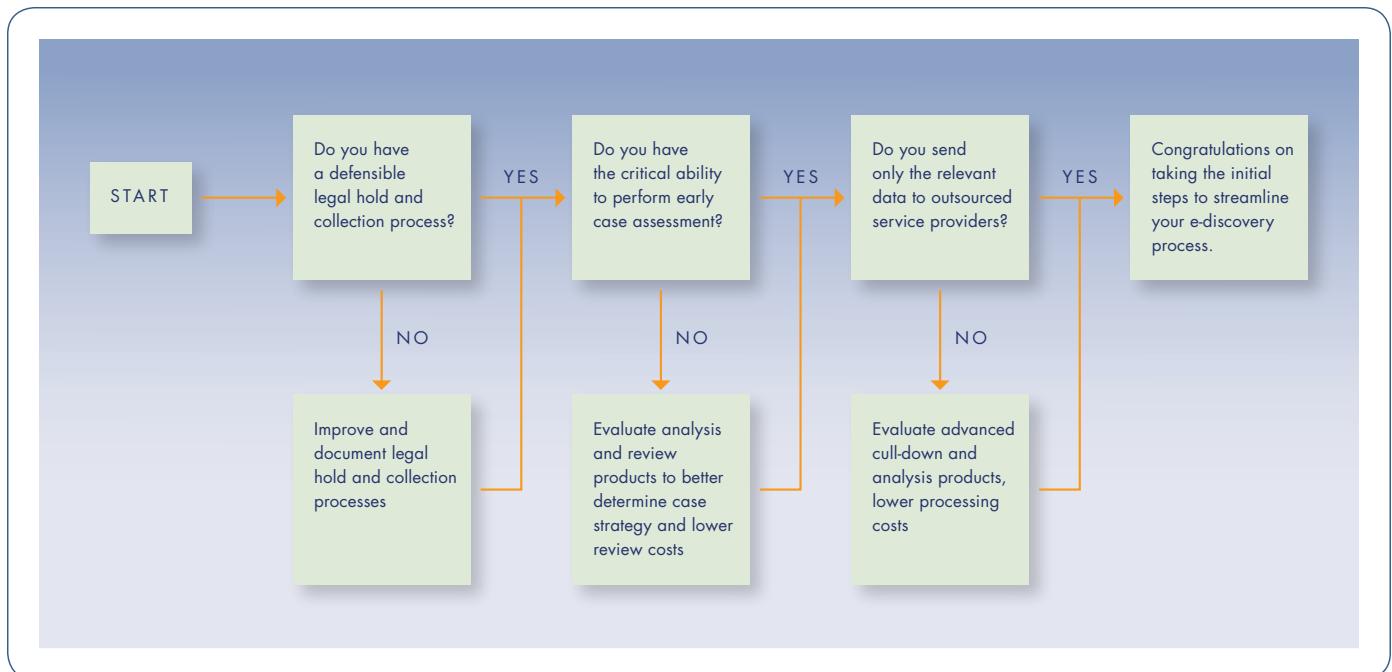


Figure 3:
E-Discovery Decision Matrix

- Define your current processes for analysis and review. Do you send all the data collected to a service provider, even though only 10-20 percent of the data is relevant to the case? If so, consider using in-house e-discovery products that can cull-down data to relevant documents before sending them to a service provider. The cost savings can be very compelling. Consider a small case with 100 GB of data: sending 20 GB of data versus 100 GB will result in cost savings of \$162,000 (80 gigabytes x \$2,000). Next-generation analysis and review products present a low-risk e-discovery improvement; they are up and running quickly, require little effort to maintain, and are also offered by e-discovery service providers.

Conclusion

Once considered an outsourced by-product of litigation, e-discovery is now viewed as a critical business process. To remain competitive and avoid staggering legal bills, companies simply have no choice but to take more control and find ways to reduce the costs, complexity, and overall risk of e-discovery. While it is easy to get caught up in the overall process and try to “boil the ocean”, it is best to focus on phases where the biggest gains can be made. But, in order to do that, e-discovery needs to be managed like a business process so that risks, costs, and overall exposure are understood and managed appropriately.

Remember that no single e-discovery vendor provides capabilities for all phases of e-discovery. Instead, focus efforts on improving your e-discovery process one step at a time and prioritize the steps based on their duration, cost, and impact to your organization. The decision matrix in figure 3 will help determine where to start. Most companies find that by first focusing on the processing, analysis, and review steps in the short-term, they are able to significantly reduce both the cost and duration of their e-discovery process.

E-DISCOVERY RESOURCES**2007 Socha-Gelmann Electronic Discovery Survey**

<http://www.law.com/jsp/legaltechnology/pubArticleLT.jsp?id=1185959203368>

The fifth annual report on electronic discovery by George Socha and Tom Gelmann.

E-Discovery 2.0

<http://www.clearwellsystems.com/e-discovery-blog/>

Thoughts by the e-discovery team at Clearwell Systems, on the evolution of e-discovery.

E-Discovery Roadmap

<http://www.law.com/jsp/legaltechnology/eDiscoveryRoadmap.jsp>

An interactive learning guide covering the major components of e-discovery.

EDDix LLC

http://www.eddixllc.com/member_home.asp

Independent research, analysis and reporting on the Electronic Data Discovery (EDD) marketplace.

Electronic Discovery Reference Model (EDRM) Project

<http://www.edrm.net>.

Launched in May 2005, the EDRM Project addresses the lack of standards and guidelines in the electronic discovery market.

The Sedona Conference

<http://www.thesedonaconference.org/>

Nonprofit research and educational institute dedicated to the advanced study of law and policy.

TechnoLawyer Problem/Solution Guide

<http://blog.technolawyer.com/2007/07/blawgworld-2007.html>

Contains 185 problems and corresponding solutions related to e-discovery.

Third Annual Litigation Trends Survey Findings

<http://www.fulbright.com/mediaroom/files/2006/FulbrightsThirdAnnualLitigation-TrendsSurveyFindings.pdf>

Fulbright and Jaworski's annual survey of litigation trends in the U.S. and the world.

FOR MORE INFORMATION

For more information about
Clearwell Systems Inc., or the
Clearwell E-Discovery Platform,
please contact us at:

Clearwell Systems

441 Logue Avenue
Mountain View, CA 94043
650.526.0600 tel
650.526.0699 fax
www.clearwellsystems.com
info@clearwellsystems.com



Clearwell