

Best Practices from Real-World Cases

# Early Case Assessment in eDiscovery

Who should read this paper

Lawsuits, corporate investigations, and regulatory audits are increasing in number—rapidly. As a result, companies are struggling to control skyrocketing eDiscovery costs. This paper details eDiscovery best practices derived from real-world electronic discovery case studies and explains how Clearwell customers have met their eDiscovery challenges and achieved a significant return on investment (ROI)—often within a single case.



# Early Case Assessment in eDiscovery

## Best Practices from Real-World Cases

### Contents

<b>Introduction</b> .....	<b>1</b>
<b>eDiscovery Best Practices</b> .....	<b>1</b>
PHASE 1: PREPARING FOR THE CASE .....	1
PHASE 2: EARLY CASE ASSESSMENT .....	4
PHASE 3: SEARCH AND CULL-DOWN .....	6
PHASE 4: REVIEW AND PRODUCTION .....	8
<b>SUMMARY</b> .....	<b>11</b>

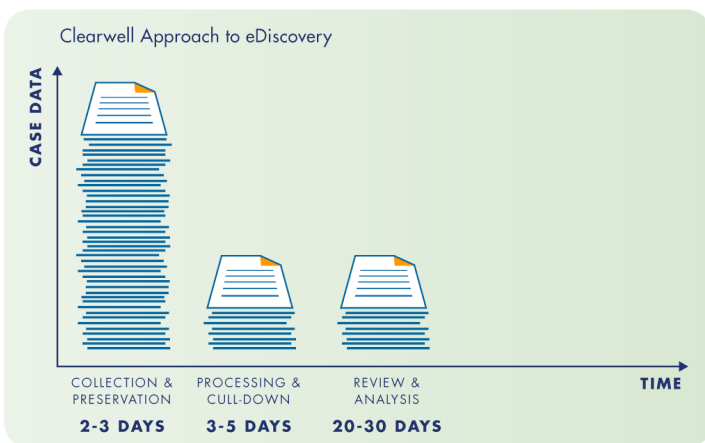
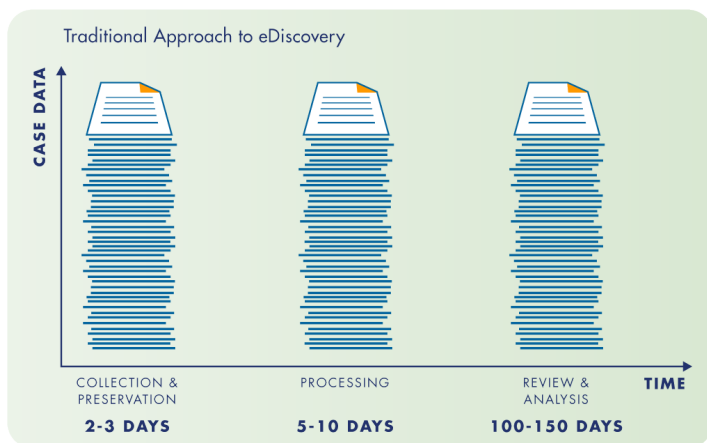
## Introduction

Lawsuits, corporate investigations, and regulatory audits are increasing in number—rapidly. In 2010, 26 percent of corporations noted that they were expecting a significant increase in litigation.<sup>1</sup> The amount of electronically stored information is also increasing. More than 90 percent of new business records are created electronically, and 40 percent of them are never converted to paper.<sup>2</sup> The net result: companies are struggling to control skyrocketing eDiscovery costs and to complete investigations with existing resources.

The traditional approach of using multiple, disparate tools to conduct eDiscovery, spending days processing all the data and weeks conducting review is slow, expensive, and risky. Companies need a new approach that allows them to address the eDiscovery challenge by working smarter, not harder. Companies today require a new level of visibility and control earlier in the eDiscovery process. This capability would allow companies accelerate early case assessments, intelligently cull down data, increase reviewer productivity, and ensure the defensibility of their eDiscovery process. Numerous Fortune 500 companies, law firms, and government agencies have benefited from faster decision making and significant cost reductions by centrally managing all of their cases using the Clearwell eDiscovery Platform®. This paper details several best practices derived from real-world cases and explains how Clearwell customers have met their eDiscovery challenges and achieved a significant return on investment (ROI)—often within a single case.

### Clearwell Benefits

- Accelerate early case assessments
- Cull down data by up to 90 percent
- Perform faster review
- Enhance the defensibility of eDiscovery
- Reduce the number of eDiscovery tools



## eDiscovery Best Practices

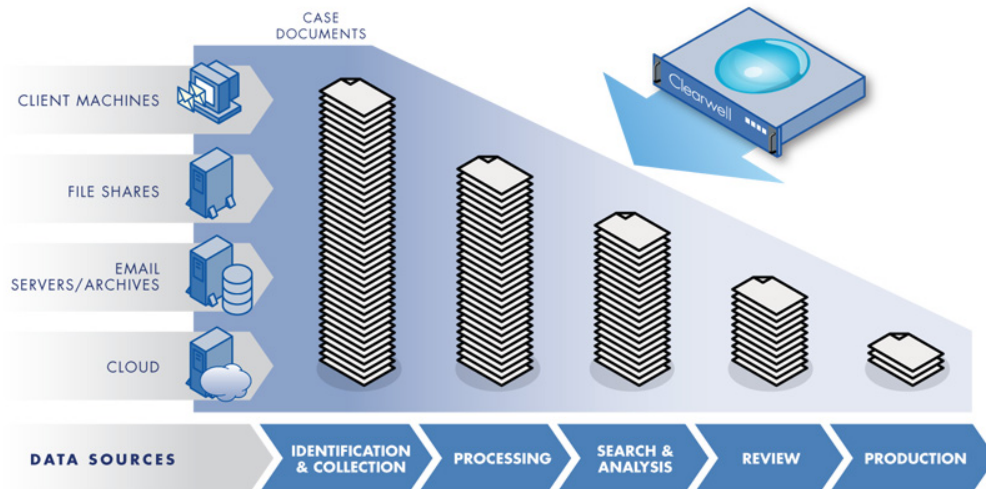
### PHASE 1: PREPARING FOR THE CASE

The first steps in any case are to audit data collections, estimate case budget and timeline, eliminate useless files, and then process the resulting data set. Traditional eDiscovery solutions take weeks to deploy and force users to wait until all full document processing is complete before they can begin analyzing a case. This processing step is often slow and takes days to complete—even for a case that only involves a small number of custodial documents.

1-2010 Fulbright and Jaworski Survey available at <http://www.fulbright.com/>  
2-Brian Ingram, "Locate Smoking Guns Electronically." COM.

## Early Case Assessment in eDiscovery Best Practices from Real-World Cases

By contrast, Clearwell is up and running within 25 minutes and integrates seamlessly with your existing eDiscovery collection process, allowing customers to begin their case processing and analysis immediately. Clearwell starts indexing documents as soon as you collect and preserve them on dedicated NAS or file shares using existing methodologies. If required, Clearwell can also index and analyze data on live data sources (for example, archives, Microsoft® Exchange, Lotus Notes®, etc.). The following key features are utilized by Clearwell customers to prepare for a case:

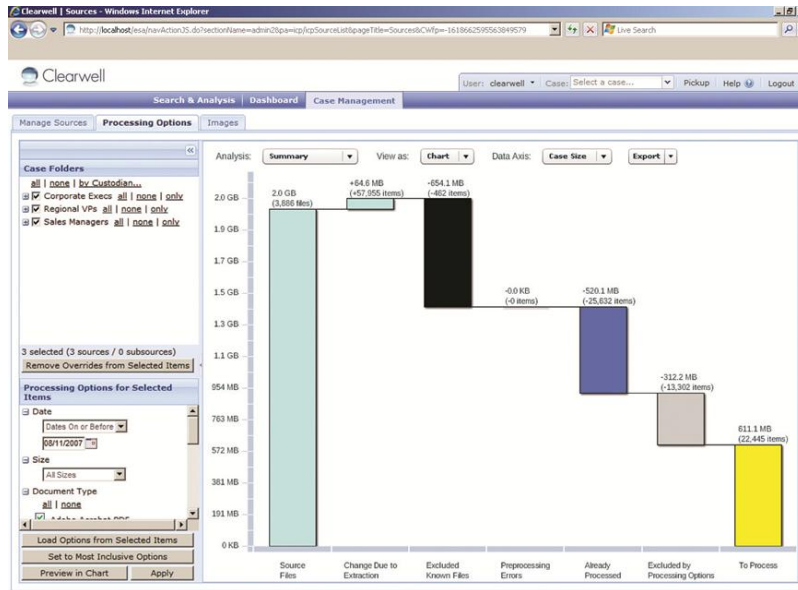


**Up and running within 25 minutes, Clearwell provides a single, integrated processing, search, analysis, review, and production platform.**

**Pre-Processing Analytics:** Prior to full document processing, Clearwell visually summarizes overall document set characteristics and presents detailed analysis by custodian, timeline, and file type. Users can quickly confirm that all case data has been collected and can use the detailed statistics to accurately estimate eDiscovery budget and timeline. This information is critical when preparing for the Meet and Confer.

**Advanced Pre-Processing Filters:** Clearwell enables users to interactively filter data by custodian, date, strong file type, and file size prior to processing. Clearwell also provides one-click filtering of data by NSRL database list (commonly known as NIST list) and custom known file lists, significantly reducing downstream processing and review costs.

## Early Case Assessment in eDiscovery Best Practices from Real-World Cases



**Pre-Processing Filters can reduce case data sets by up to 30 percent by filtering on custodian, date, strong file type, and file size prior to processing.**

**Rapid Processing of Emails, Attachments, and Documents:** Clearwell's processing is 3 to 5 times faster than other eDiscovery products, allowing users to process even large cases in a matter of hours. Unlike other solutions that require customers to wait until processing is complete, Clearwell's multi-threaded architecture enables customers to begin their case assessments immediately. During processing, Clearwell automatically deduplicates data with customers typically experience deduplication rates between 30 and 60 percent.

**Multi-Language Support:** Clearwell provides full Unicode compliance and supports English, Western European, Eastern European, Cyrillic, and Asian languages. Clearwell also enables automatic language identification of documents and provides exact document counts by language type across the entire data set.

**Incremental Processing:** Clearwell allows customers to process additional case documents at any time without re-creating the entire case index. This saves significant time and resources because the discovery of new information or additional custodians occurs frequently during most cases. Clearwell can also be easily configured to process only new information from a data source on a scheduled and reoccurring basis to ensure the index is always current.

### Real-World Use Cases

A technology company used Clearwell for eDiscovery in response to a lawsuit filed by an employee who was allegedly discriminated against. The lawsuit cited proof within the corporate email system. In this case, the customer installed Clearwell, and within 25 minutes they had pre-processed two large custodial PST files and 14 GBs of files from network directories, and audited the collected data using graphical timeline analytics. They noticed two months of emails were missing and contacted IT to collect the missing data to ensure that no data "time-gaps" existed and avoid potential sanctions during trial. Next they de-NISTed the collection to remove non-user generated system files and other multimedia files that were not relevant to the case and eliminated 40 percent of the data before processing. Using pre-processing analytics, the customer evaluated the potential time and cost requirement for processing and reviewing the data collection. They proceeded to process the collection and began analysis within minutes. With key evidence discovered in a matter of hours, the legal team had more time to weigh legal options and choose the right case strategy, that is, to settle or fight the lawsuit in court.

A large insurance company used Clearwell for eDiscovery in response to a regulatory inquiry. They needed to analyze email from live Microsoft Exchange servers (which contained the most recent six months of data) and email from their old server backup tapes. The first step in the process was to give Clearwell access to their Microsoft Active Directory® server to collect organizational information. Within 10 minutes, Clearwell had a list of all Exchange servers around the world, all mailboxes within each server, and other information such as end-user aliases, department, location, and distribution lists. Then, the customer simply started the Clearwell indexing process by selecting all servers and a specific date range. Next, the customer converted several backup tapes to PST files and placed them onto a network share. However, the date ranges of the PST files overlapped with that of the live Exchange servers. This redundant data was removed by giving Clearwell access to the network share to create a single, deduplicated master index. Clearwell eliminated 60 percent of the data as part of its deduplication process. With all of the case data in a single index, the customer used Clearwell's quick search interface to perform the required searching and tagged all the responsive data into an organized project to be viewed by the regulator. Their objective was to be "responsive without being overly inclusive." The customer then created a secure account for the regulator. This allowed the regulator to log into Clearwell (via the Clearwell Web-based interface) and review only the responsive data. By using Clearwell to respond to this regulatory inquiry, the customer shortened the time needed to obtain, index, and analyze the data by several weeks.

### **PHASE 2: EARLY CASE ASSESSMENT**

Once processing is complete, the traditional approach to eDiscovery has forced customers and their law firms to launch a massive "review effort" that entails hiring several contract attorneys who are allocated portions of the data for analysis. These reviewers spend many days or weeks sifting through false positives, irrelevant messages, and redundant data. As a result, it is often weeks before the true context of the case and case strategy are known.

With Clearwell, early case assessment is completed well before manual review. Inside and outside counsel teams use Clearwell's Web-based interface to rapidly navigate through discussion threads, custodial documents, message attachments, and more to get a better sense of their position on each case. This new approach from Clearwell is about working smarter, not working harder. The Clearwell approach enables early case assessment and "data cull-down" (Phase 3, discussed later) to be performed much earlier in the eDiscovery process, which leads to significant productivity gains, more deadlines met, and superior case knowledge.

With Clearwell, companies and law firms get fast answers to questions such as:

- Can we quickly find a smoking gun and locate all of its instances?
- Can we determine who knew what and when?
- How do I know what to look for?
- How do I identify the best search terms in preparation for the Meet and Confer?
- How do we find all the email address and domain permutations for certain individuals?
- Can we quickly determine how many documents are responsive and, thus, if we can meet the deadline?
- Are we sure that we have identified all the custodians and all the data relevant to the case?

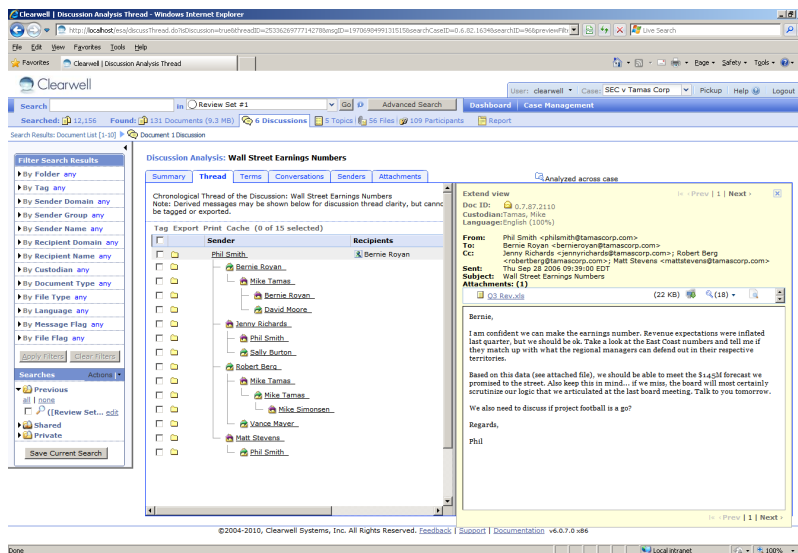
Answers to these questions help companies determine the critically important early case assessment. By knowing early in the process whether they should settle a case or prepare for court, companies realize significant cost savings. Clearwell helps customers answer these

## Early Case Assessment in eDiscovery Best Practices from Real-World Cases

early case assessment questions quickly and accurately, well before the lengthy manual review process. Early case assessment features from Clearwell include:

**Quick Search:** Clearwell's quick search capability delivers Google-like search capabilities to the entire corpus for a case, allowing customers to achieve a "first look" at a case and perform rapid early case assessment. Clearwell's search results are typically displayed in less than 10 seconds, even when searching millions of documents.

**Relevance Rank:** Clearwell's patent-pending algorithms consider the unique properties of email and documents to display the most relevant documents or discussion threads first, helping customers to assess and analyze cases more quickly. Customers often relate Clearwell's Relevance Rank technology to Google's PageRank™ technology (which delivers relevance-ranked search results from millions of web pages).



**Discussion Threads automatically link emails together to quickly identify all the custodians and determine exactly who knew what and when.**

**Discussion Threads:** Clearwell's patent-pending algorithms dynamically link together all related messages into chronological threads that capture the entire discussion, including all replies, carbon copies, and forwards. By walking the thread, Clearwell guides customers to new evidence, quickly identifies all the participants, and determines exactly who knew what and when. This allows customers to determine who took part in a particular email conversation and discover if additional custodians should be added to the case.

**Custodian Permutation Analysis:** Clearwell discovers permutations for each target custodian. This capability automates the manual process that litigation teams must perform to determine all the combinations of a custodian's name, their various email addresses, and the various domains from which they are communicating.

**Topic Clustering:** Clearwell's patent-pending linguistic algorithms automatically organize documents into specific topics, enabling customers to quickly analyze all documents related to a particular topic and identify secret project names and code words.

**File Analytics:** Clearwell identifies duplicate files that may be attached to multiple emails or may be "loose" on a user's hard drive. File analytics allow investigators to easily determine everyone who possesses or has sent or received a file of interest and allow reviewers to review a file once instead of multiple times.

**People Analytics:** Clearwell analyzes individual and group-to-group communications within a company, or to their customers, suppliers, and partners. This allows customers to easily access a list of top custodians for a search or monitor communications between regulated and non-regulated divisions within their company.

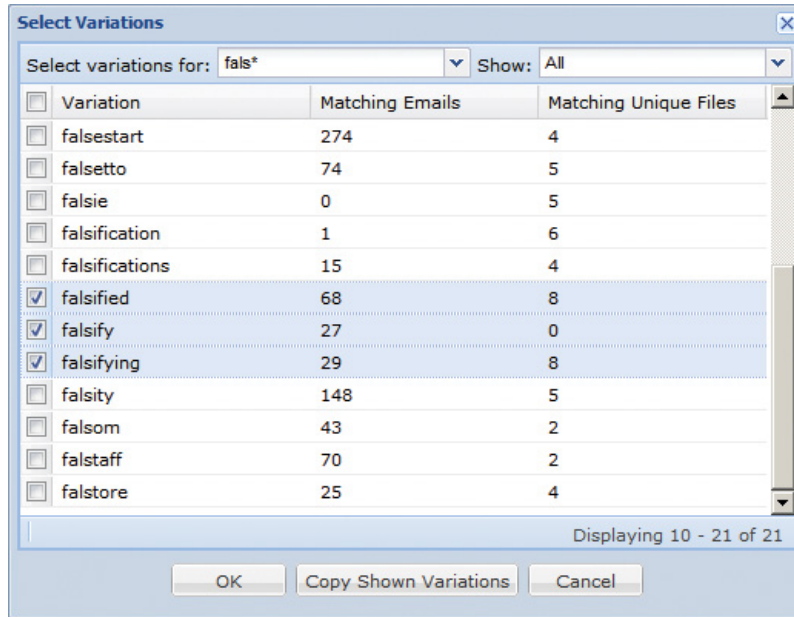
### Real-World Use Case

The forensics and litigation support teams at a Fortune 10 company use Clearwell to streamline how they handle internal investigations (for example, employee harassment, inappropriate use of company resources, wrongful termination, and intellectual property theft). In one recent investigation, several key employees abruptly left the company at the same time. Combining this with the fact that the quality of their work was also suspect just prior to departure, management quickly collected their NSF files and desktop files and indexed them using Clearwell in order to determine if there was any violation of corporate policy. However, they did not know where to begin the investigation. By using Clearwell's ability to order email discussion threads by the number of messages, they began the investigation by reading the longest conversations first. From within these conversations, the client used Clearwell's Terms feature, where Clearwell's natural language algorithms suggested the most frequent terms (that is, noun phrases) used in the conversation. Surprisingly, the most common term was "Project Escape." Within minutes, the investigation had taken a new course. Project Escape was the code word used by the employees to refer to the collection of proprietary information before their departure. The customer then searched for "Project Escape" and reviewed all the discussions that referenced this code word. Within seconds, Clearwell's Discussion Thread capability identified additional individuals who were part of the secret plan, but had not yet resigned. The customer was able to quickly remedy the situation before any material harm occurred. In the customer's own words, "We were able to accomplish in one hour with Clearwell, what would have taken us days with our former process."

### PHASE 3: SEARCH AND CULL-DOWN

In this phase, customers reduce a large dataset to a much smaller, more relevant data set. Many organizations simply process all collected data and cull-down by date range, keywords, and hit counts. Because this cull-down step reduces data sets only by an average of 30 percent to 40 percent, organizations pay a tremendous price including steep fees to outside parties for processing and legal review. Additionally, organizations are left open to sanctions should eDiscovery defensibility come into question. But with Clearwell, customers quickly and accurately cull-down large data sets by 80 percent to 90 percent and no longer need to send irrelevant data for processing and review, saving a considerable amount of time and money. Furthermore, Clearwell provides defensibility by adding transparency and automatic reporting to the eDiscovery search process, thus lowering the risk of sanctions and the loss of attorney-client privilege and work product protection. Key Clearwell features that provide search, cull-down, and defensibility include:

**Transparent Search:** Clearwell provides a set of search features that enable a more defensible and collaborative eDiscovery search process and enhance the ability to cull irrelevant information. These include:



**Search Preview provides the list of keyword variations matching a query, allowing users to selectively include and exclude keyword variations.**

- **Search Preview:** Clearwell provides visibility into matching keyword variations for wildcard and stemming searches prior to running a search. Customers can selectively include relevant variations or exclude false positive variations in their search query, removing irrelevant documents from search results.
- **Search Filters:** In real-time, customers can filter their result for individual queries or variations, and sample the filtered documents to evaluate the effectiveness of their search. This also allows customers to rapidly identify false positive documents prior to review.
- **Search Report:** Clearwell creates a comprehensive report that documents all search criteria and provides detailed analytics of the results for both the overall search and the individual queries within the search. The report tracks search terms that were included and excluded during search preview, providing a defensible audit trail of search refinement decisions.

**Multiple Query Search:** Clearwell delivers the ability to run large numbers of queries simultaneously and provides reporting for both the overall search and the individual queries within the search. Large numbers of queries can be tested in minutes not days, dramatically decreasing the turnaround time needed to evaluate the effectiveness of keyword searches.

**Auto-Filters:** With the simple click of a checkbox, Clearwell immediately excludes certain domains, locations, groups, document types and end-users from the entire set of search results—even results containing millions of documents. This makes it very easy to eliminate attorney-client privileged and irrelevant information such as lunch menus, spam, disclaimer text, and newsgroup emails.

**Communication Flow Tracking:** Clearwell further enables customers to cull down data by limiting search results based on email flow: individual-to-individual, individual-to-group, and group-to-group, including internal or external groups. With this powerful feature, Clearwell customers have reduced their case data sets by as much as 70 percent in some cases.

**Advanced Search:** Clearwell goes beyond basic keyword search and provides users with the ability to construct complex searches based on senders, recipients, communication flows, direction, subject, attachment, date range, tag values, and comments. Clearwell supports both

stemmed (that is, a search for the word "run" would return "ran," "running," "runner," etc.) and literal searches, and also provides power-user capabilities including Boolean, wildcard, fuzzy, and proximity searches.

**Nested Search:** A noteworthy differentiator, Nested Search allows all of these searches to be executed multiple times on previous search results, thereby significantly culling down large case data sets.

### Real-World Use Cases

A manufacturing customer wanted to investigate conversations between a set of employees and several outside entities. They collected desktop files and PST files containing over 500,000 emails and loose file documents, but were running out of review time with only three days left to respond to opposing counsel. Faced with a difficult situation, Clearwell was brought in to help. As promised, Clearwell was installed and completed the indexing process in less than four hours. Clearwell's automatic deduplication process immediately reduced the emails and documents by 20 percent, to 400,000. Next, the customer quickly excluded attorney-client privileged emails by selecting all messages sent to and from their outside counsel's domain, which reduced the data set to 360,000 emails and documents. Then, the customer performed a multiple query keyword search and tagged the resulting 250,000 emails and documents into a responsive dataset. From here, the customer further filtered the responsive data set to emails and documents sent between five employees and two external client domains, significantly reducing the number of emails and documents to 15,000. Using Clearwell's Web-based interface, the customer evenly allocated the 15,000 emails and documents to separate remote reviewers for review. Using role-based security, the customer easily ensured that the reviewers only had visibility to their respective data sets (described in more detail in Phase 4 below). The end result: the customer reduced the data that needed to be reviewed by 97 percent (from 500,000 emails and documents to 15,000 responsive emails and documents) in less than eight hours and met their deadline with time to spare.

A large financial services company sent case data to an outside eDiscovery service provider to process and load into a litigation support tool at \$2,000 per GB (for the purpose of Bates stamping, TIFF conversion, and redaction). With the growing number of cases, their costs were spiraling out of control. In an effort to control costs, they purchased Clearwell to cull down the data set to the most relevant documents prior to sending it to the service provider for processing. In their first case, they were able to cull down the data set from 100 GB to approximately 20 GB in three days—an 80 percent reduction. Clearwell saved the company from producing 80 GB of data and spending \$160,000 (80 GB x \$2,000 = \$160,000), delivering a 200 percent ROI from this single case. In addition, Clearwell provided a documented audit trail of all searches, helping the company demonstrate care and reasonableness in their eDiscovery search and cull-down process when questioned by opposing counsel in court.

### PHASE 4: REVIEW AND PRODUCTION

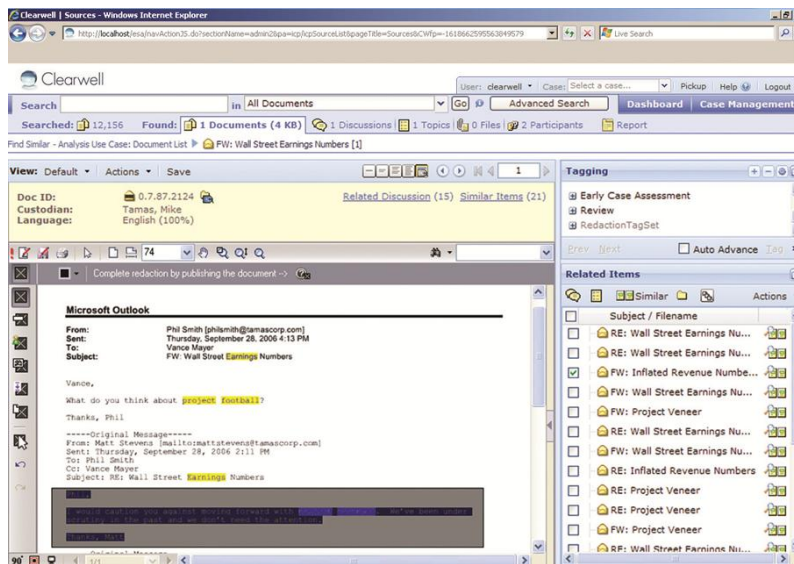
The next phase in the eDiscovery process is review and production. The traditional approaches consist of: (1) printing each document and manually reviewing them with highlighters in hand, or (2) converting them into TIFF and reviewing in a litigation database tool. Both approaches are fraught with high costs, limitations, and errors. Further, they do not provide the much-needed context to perform an accurate examination of the document, and fail to uncover important metadata such as formulas and hidden columns in a spreadsheet for example. Clearwell delivers several new capabilities that increase the efficiency and consistency of the review process. Clearwell presents each document in complete context, displaying related documents, such as those contained in discussion threads and attachments, improving review consistency, eliminating redundant review, and enabling bulk decisions. Customers can eliminate the unnecessary conversion to TIFF, quickly view all the data in native format with an integrated Native Viewer, redact sensitive information, and ensure consistency across reviewers. Clearwell displays documents exactly as they will be produced, and provides the option to produce in native, TIFF, or PDF after review is completed. For review and production, Clearwell customers typically:

## Early Case Assessment in eDiscovery Best Practices from Real-World Cases

1. Review the responsive data within Clearwell.
2. Auto-redact specific keywords, phrases, and personal information.
3. Redact sensitive work-product or attorney-client information and apply redaction reason codes.
4. Produce or export the relevant document set.

Key features for the efficient review of responsive documents using Clearwell include:

**100 Percent Web-based Review:** Business professionals in legal, compliance, and human resources, for example, can now have secure access to review documents assigned to their projects using a standard Web browser. Clearwell's role-based security enables case administrators to grant access to just a single case or set of cases, enabling secure access for any number of internal constituents as well as external counsels. Web-based access and review eliminates the need to physically transport evidence, and significantly reduces the costly "back-and-forth" between teams when additional analysis needs to be done.



**Native Viewer** provides immediate access to documents, attachments, and email threads in native formats, enabling rapid and collaborative review, tagging, and redaction.

**Native Viewer:** Clearwell provides the ability to view documents in their native format without requiring each application to be loaded on a reviewer's workstation. Both text search and hit highlighting within documents are also available, increasing reviewer productivity.

**Redaction:** Clearwell enables reviewers to quickly and easily redact documents in multiple colors, apply reason codes, and verify redactions prior to production. Reviewers can redact specific text, pages, or areas within a document.

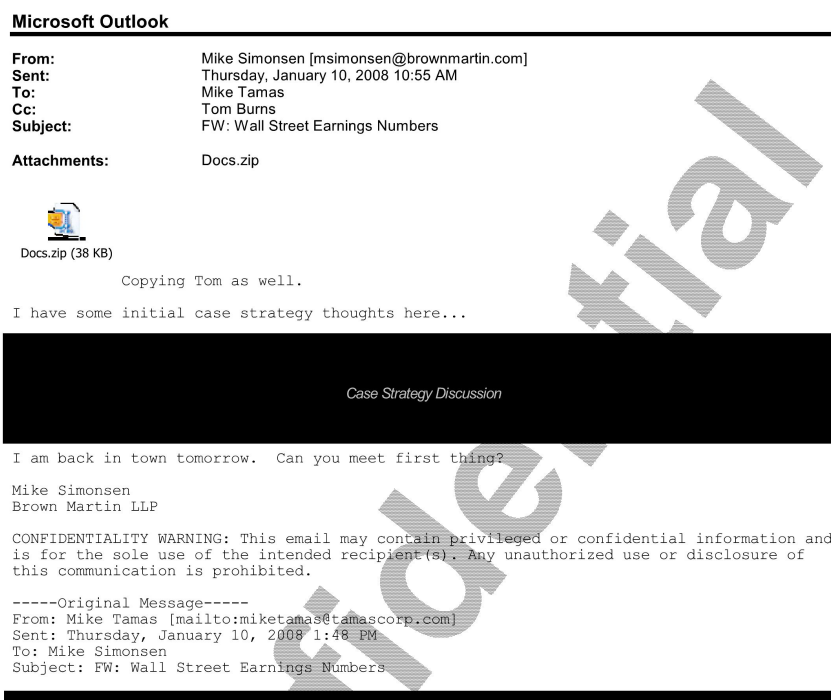
**Decision-Tree Tagging:** Clearwell creates the ability to tag individual or sets of documents using a multi-layer tree structure. This directs reviewers into key decision points, preventing errors and minimizing the number of clicks needed to accurately tag a document. Customization capabilities allow administrators to create different tag sets and decision trees and control tag access.

## Early Case Assessment in eDiscovery Best Practices from Real-World Cases

**Related Items Window:** Clearwell automatically identifies and displays related documents, such as those contained in discussion threads and attachments, to the document being reviewed. Reviewing related documents together improves consistency, eliminates redundant review, and enables bulk decisions.

**Persistent Hit Highlighting:** Clearwell highlights search terms in messages, attachments, and files within the Native Viewer, allowing reviewers to simultaneously view highlighted search terms from any number of previously performed searches (for example, terms from previously executed privilege and responsive searches).

**Case Analytics:** Clearwell delivers summary metrics on the status of all cases. Charts display the total number of messages reviewed, un-reviewed, and tagged into specific categories. Pre-built reports automatically monitor progress to ensure that resources are properly allocated and deadlines are met.



**Clearwell provides document productions in TFF format branded with redactions, Bates stamps, and other custom labels.**

Key features for production or export of responsive data include:

**Native and Image-Based Production:** Clearwell allows users to produce documents in native, TIFF, or PDF formats. Flexible export options allow for multiple metadata formats (EDRM XML, DAT, CSV, etc.), "reduping" of data by custodian, and reservation of original folder hierarchies.

**Bates Stamping:** Clearwell automates the sequential numbering of documents with Bates stamps and supports the application of custom labeling of produced documents using headers, footers, and watermarks.

**Rolling Production:** Clearwell enables batch productions at any point in the eDiscovery process. Users can track the progress of each production in real-time and view produced document sets within Clearwell in the exact format provided to outside parties.

### **Real-World Use Case**

A law firm used Clearwell for eDiscovery on a large manufacturing client's financial fraud investigation. In this case, the client employed the law firm to perform the review of 50 NSF files containing over 190,000 messages and 50 desktop file collections containing 95,000 documents. The deadline imposed by the regulators to produce only the relevant, non-privileged information was extremely aggressive—less than two months. In addition, the law firm was advised by the client to complete this task within a tight budget. Once the data was indexed by Clearwell, the case administrator used Clearwell's bulk-tagging capabilities to divide the emails and documents between eight lawyers (who were physically located throughout the U.S.), and gave each secure, Web-based access to Clearwell. The case administrator used role-based security to ensure that each lawyer only had access to their assigned data set. Immediately, Clearwell saved enormous amounts of time and money since the law firm could easily leverage a larger pool of resources without costly travel or physically sending large amounts of data. The lawyers proceeded to efficiently review and tag emails and attached documents by discussion threads, and then by individual messages. During the case, a new set of keywords was added. Because Clearwell's review and production modules are integrated with the search and analysis functionality, the client was able to iterate between phases and quickly perform additional searches and add responsive emails and documents to the review set in real time without exporting and importing between tools. The case administrator used Clearwell's case analytics capability to provide daily reports to law firm partners and the client, keeping them updated on the progress throughout each step of the review process. Once complete, the case administrator previewed and produced only the responsive emails and documents in TIFF format and delivered them to opposing counsel. The firm delivered accurate results on time and on budget for their client—a feat that wasn't possible prior to Clearwell.

### **SUMMARY**

Today's business environment has fueled a significant increase in the number of lawsuits, corporate investigations, and regulatory inquiries. As a result of this rapid growth, it is no surprise that eDiscovery costs are spiraling out of control. Companies are under immense pressure to contain costs, more efficiently use internal resources, meet court-appointed deadlines, and defend their eDiscovery process. Companies are reducing this pressure by adopting a new approach to eDiscovery. By using Clearwell to process, analyze, search and cull, and review and produces case data sets, companies are transforming their eDiscovery processes. The end result: corporations and law firms accelerate early case assessments from weeks to hours; cull down data by up to 90 percent; increase review throughput and consistency; reduce the movement of data across multiple, disparate tools; and improve the defensibility of their eDiscovery process.



## About Symantec

Symantec is a global leader in providing security, storage, and systems management solutions to help consumers and organizations secure and manage their information-driven world. Our software and services protect against more risks at more points, more completely and efficiently, enabling confidence wherever information is used or stored. Headquartered in Mountain View, Calif., Symantec has operations in 40 countries. More information is available at [www.symantec.com](http://www.symantec.com).

For specific country offices and contact numbers, please visit our website.

Symantec World Headquarters  
350 Ellis St.  
Mountain View, CA 94043 USA  
+1 (650) 527 8000  
1 (800) 721 3934  
[www.symantec.com](http://www.symantec.com)

Symantec helps organizations secure and manage their information-driven world with **IT Compliance, discovery and retention management, data loss prevention, and messaging security** solutions.

Copyright © 2011 Symantec Corporation. All rights reserved. Symantec, the Symantec Logo, and the Checkmark Logo are trademarks or registered trademarks of Symantec Corporation or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.  
7/2011 21198622