Clearwell

E-DISCOVERY BEST PRACTICES
from Real-World Cases

A Clearwell White Paper

Table of Contents

Introduction 3

E-Discovery Best Practices 4

Phase 1: Preparing for the Case 4

Phase 2: Early Case Assessment 5

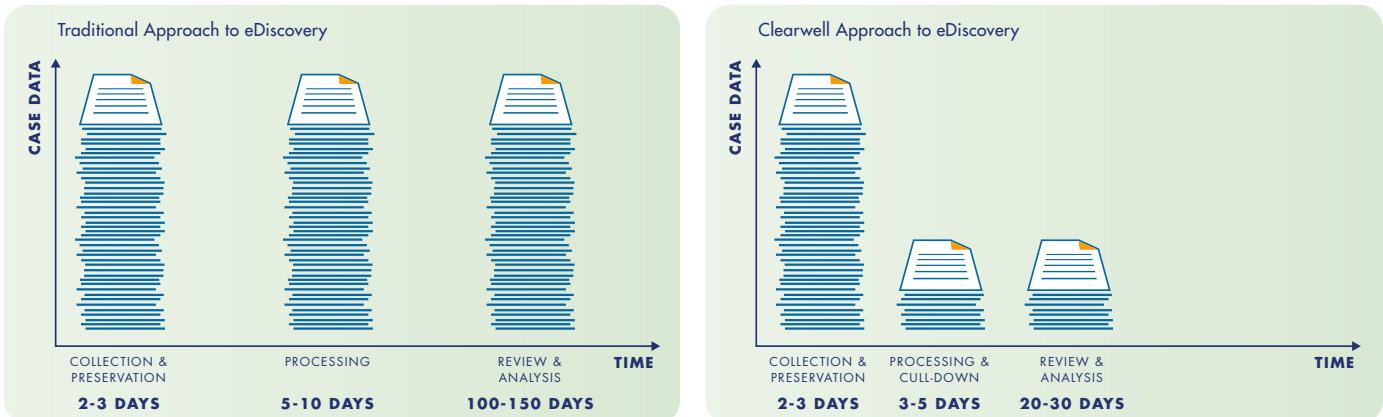
Phase 3: Search and Cull-Down 7

Phase 4: Review and Export 8

Summary 10

Introduction

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% from the previous year.1 The amount of electronically stored information is also increasing. More than 90% of new business records are created electronically, and 40% of them are never converted to paper.2 The net result: companies are struggling to control skyrocketing e-discovery costs and to complete investigations with existing resources. The traditional approach of 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 e-discovery challenge by working smarter, not harder. Companies today require a new level of visibility and control



Clearwell Benefits

- Perform early case assessments in hours vs. days
- Lower processing costs by up to 80%
- Reduce review workload by up to 90%

earlier in the e-discovery process. This capability would allow companies to perform rapid early case assessments, lower processing costs, reduce review workload, and regain control of their e-discovery process. Numerous Fortune 500 companies, law firms, and governmental agencies have benefited from faster decision making and significant cost reductions by centrally managing all of their cases using the Clearwell E-Discovery Platform® (Clearwell). This paper details several best practices derived from real-world cases and explains how Clearwell customers have met their e-discovery challenges and achieved a significant return on investment (ROI)—often within a single case.

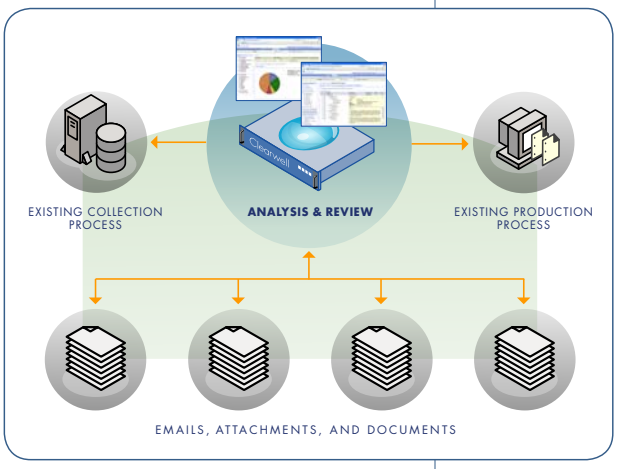
.

E-Discovery Best Practices

PHASE 1: PREPARING FOR THE CASE

The first step in any case is to process the data to be analyzed. Traditional e-discovery solutions take 1-3 days to deploy and force users to wait until all processing is complete before they can begin analyzing a case. This processing step is often slow and takes hours to complete—even for a case that only involves a small number of custodial documents.

By contrast, Clearwell is up and running within 25 minutes and allows customers to begin their case 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 (e.g. archives, Microsoft® Exchange, etc.). The following key features are utilized by Clearwell customers to prepare for a case:



Clearwell analyzes emails, attachments, and documents without requiring any changes to your current collection or litigation production processes.

Rapid processing of emails, attachments, and documents: Clearwell's processing is fast and efficient. While actual processing speeds vary based on the average message size, Clearwell processes 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.

Organizational Discovery: Clearwell discovers end-user email aliases and domain information, and automatically creates filters by domain, geography, department, and end-user. This enables customers to immediately weed out false positives, irrelevant files, and privileged information with the click of a mouse. Customers can also configure Clearwell to discover and analyze Microsoft Active Directory® information such as distribution lists, department, and location information.

De-duplication across Data Sources & Custodians: During processing, Clearwell's algorithms identify duplicate documents, and treat them as a single entity, eliminating the subsequent review of the same content by multiple people. De-duplication rates vary based on the type and number of data stores indexed. Clearwell's customers typically experience de-duplication rates between 30-60%.

Incremental Processing: Clearwell allows customers to process additional case documents at any time without recreating the entire case index. This saves significant time and resources since 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 e-discovery in response to a lawsuit filed by an employee who was allegedly discriminated against. The lawsuit sited proof within the corporate email system. In this case, the customer installed Clearwell, de-duplicated and indexed 2 large custodial PST files, and began their analysis in approximately 25 minutes. With key evidence discovered in a matter of minutes, 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 e-discovery in response to a regulatory inquiry. They needed to analyze email from live Microsoft Exchange servers (which contained the most recent 6 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, de-duplicated master index. Clearwell eliminated 60% of the data as part of its de-duplication 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 e-discovery 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 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.

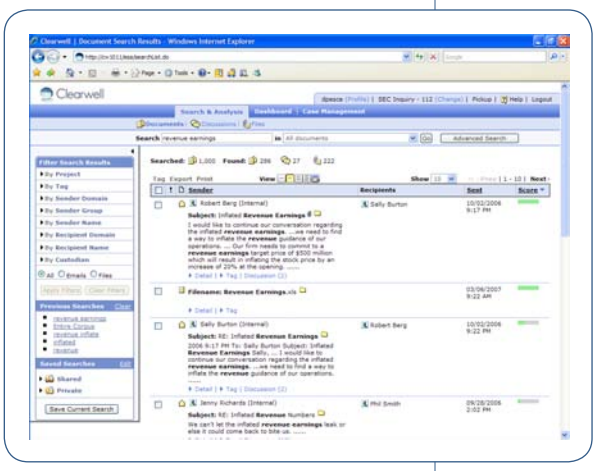
With Clearwell, early case assessment is completed well before manual review. Inside and outside counsel teams use Clearwell's web-based interface to rapidly navigate through discussion threads, custodial documents, message attachments, and more to get a better sense of their position on each case. This 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 below) to be performed much earlier in the e-discovery 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 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 questions quickly and accurately, well before the lengthy manual review process. Early case assessment features from Clearwell include:

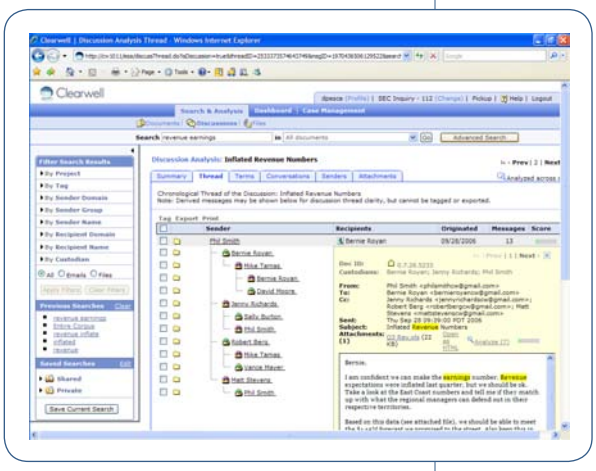
Quick Search: Clearwell's quick search capability delivers Google-like search capabilities to the entire corpus for a case, allowing customers to achieve a "first look" at a case and perform rapid early case assessment. Clearwell's search results are typically displayed in less than 10 seconds, even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



Relevance Rank speeds analysis by displaying the most relevant emails and documents first.

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: 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.



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

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 Classification: Clearwell automatically organizes data into specific topics based on patent-pending linguistic algorithms, guiding customers to find important 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.

Real-World Use Case

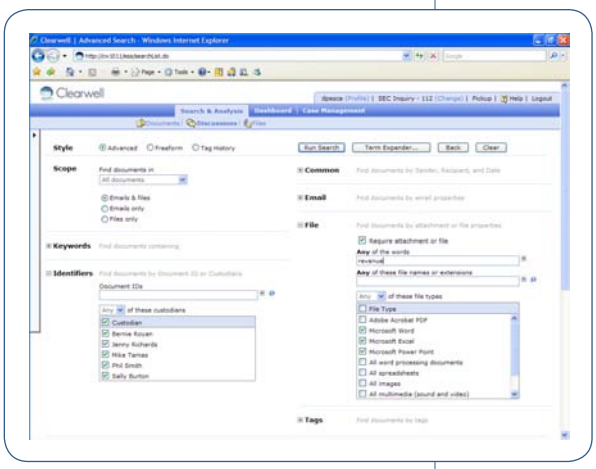
The forensics and litigation support teams at a Fortune 10 company use Clearwell to streamline how they handle internal investigations (e.g., employee harassment, inappropriate use of company resources, wrongful termination, and intellectual property theft). In one recent investigation, several key employees 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 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 (i.e., 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

The cull-down process begins in parallel or right after early case assessment. In this phase, customers reduce a large dataset to a much smaller, relevant dataset. But the traditional e-discovery process doesn't include a cull-down step. Many organizations simply process and review all the potentially relevant information and therefore pay a tremendous price including steep fees paid to outside parties for processing and legal review. But with Clearwell, customers quickly and accurately search and cull-down large datasets and no longer need to send irrelevant data for processing and review, saving a considerable amount of time and money. Key Clearwell features that drive this time and cost savings include:

Search speed: All search results are returned in less than 10 seconds even when finding hundreds of thousands of documents within cases containing millions of documents. Results can be immediately analyzed to determine if false positives are present and providing leverage to negotiate with opposing counsel on the search terms using actual numbers.



Advanced Search delivers powerful and flexible capabilities to cull-down large datasets of messages, attachments, and documents.

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, group-to-group including internal or external groups. With this powerful feature, Clearwell customers have reduced their case datasets by as much as 70% 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 (i.e., 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 datasets.

One-click Tagging: Clearwell's one-click tagging capability allows users to tag entire email discussions or complete result sets to rapidly separate the wheat from the chaff during the cull-down process.

Real-World Use Cases

A manufacturing customer wanted to investigate conversations between a set of employees and several outside entities. They collected PST files containing over 500,000 messages, but were running out of review time with only three days left to respond to opposing counsel. Faced with an impossible situation, Clearwell was brought in to help. As promised, Clearwell was installed and completed the indexing process in less than 4 hours. Clearwell's automatic de-duplication process immediately reduced the messages by 20%, to 400,000. Next, the customer quickly excluded attorney-client privileged email by selecting all messages sent to and from their outside counsel's domain, which reduced the dataset to 360,000 messages. Then, the customer performed a multi-phrase keyword search and tagged the result of 250,000 messages into a responsive dataset. From here, the customer further filtered the responsive dataset to emails between five employees and two external client domains, significantly reducing the number of messages to 15,000. Using Clearwell's web-based interface, the customer evenly allocated the 15,000 messages to separate remote reviewers for review. Using role-based security, the customer easily ensured that the reviewers only had visibility to their respective datasets (described in more detail in Phase 4 below). The end result: The customer reduced the data that needed to be reviewed by 97% (from 500,000 messages to 15,000 responsive messages) in less than eight hours and met their deadline with time to spare. A large financial services company sent case data to an outside e-discovery 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 dataset 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 dataset from 100GB to approximately 20GB in three days—an 80% reduction. Clearwell saved the company from producing 80GB of data and spending $160,000 (80GB x $2,000 = $160,000), delivering a 200% ROI from this single case.

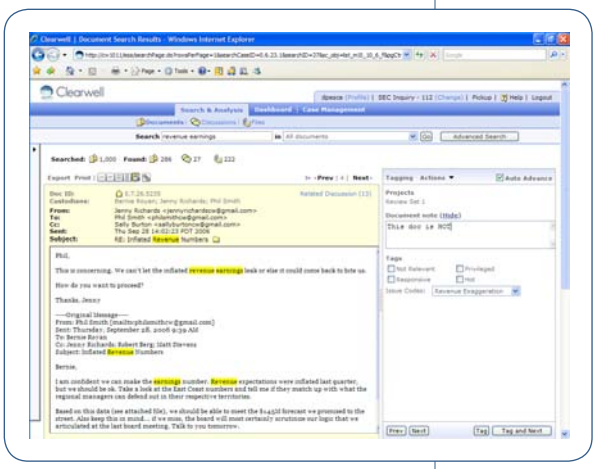
PHASE 4: REVIEW AND EXPORT

The next phase in the e-discovery process is detailed 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. Clearwell delivers several new capabilities that increase the efficiency and accuracy of the review process. Clearwell presents each document in complete context, enabling reviewers to make the most accurate decisions. Customers can eliminate the unnecessary conversion to TIFF, quickly view all the data in HTML or native format, ensure consistency across reviewers, and deliver only relevant data

to external parties. For cases that require document production, Clearwell provides easy integration with litigation production tools and services. For review and export, Clearwell customers typically:

1) Review the responsive data within Clearwell or
2) Export the responsive data and send it to the requesting party which is typically a regulator or opposing counsel, or
3) Export the responsive data for production by a litigation support company in order to prepare for court

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

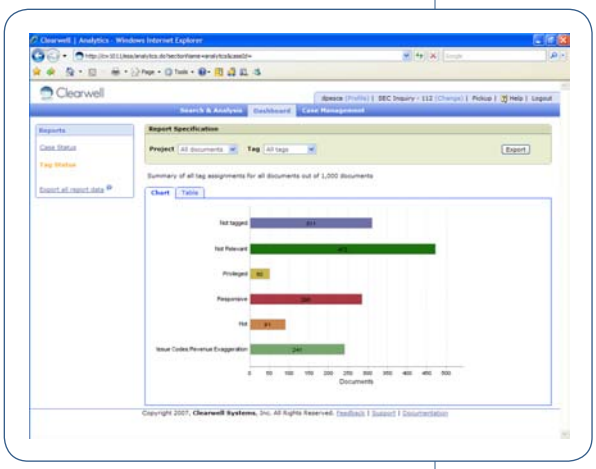


One-Click Tagging instantly tags individual emails and documents.

100% 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.

Review Modes: Clearwell provides multiple review modes giving users the flexibility to view header, snippet, or document detail. Reviewers can review documents individually, or review them grouped by context (e.g., by discussion threads, topics, etc.), increasing review throughput. Each review mode is optimized to maximize screen real estate, and minimize mouse clicks and mouse movement, further increasing review throughput.

One-Click Tagging: One-click tagging capabilities allow users to tag individual or bulk email and documents into projects, update the status as they relate to a project, and assign them to an individual for review. Users can easily configure Smart-Tags that automatically apply tags to future email and documents that meet the pre-defined criteria.



Case Analytics track the status of all cases to ensure that resources are properly allocated and deadlines are met.

Discussion Thread Tagging: Clearwell also allows customers to review and tag the entire email conversation with a single click, speeding the review process. Further, by taking action on the entire thread, customers can ensure consistency and accuracy during the review process.

Hit-highlighting: Clearwell previews search results in HTML and highlights search terms in email messages, attachments, and documents—allowing you to quickly find what you are looking for without having to read every single word. Clearwell supports over 400 different document and attachment types, including multiple word processing formats, spreadsheet formats, presentation formats, emails as attachments, and compressed formats such as .zip.

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.

Privilege log: Clearwell automatically creates a "privilege log" report of all privileged documents for opposing counsel's review. Users have the flexibility to select the fields

to include in the report, such as document ID, custodian, to, from, subject, title, and reason codes and produce in a CSV format enabling additional post-processing if needed.

Key features for exporting the responsive data include:

Native, PDF and XML Export: Clearwell allows results to be exported in native formats, via EDRM compliant XML or converted to Adobe PDF® format. Users can re-dupe data by custodian during export, password protect results, preserve original folder hierarchy, and control the number of files created (one file for all custodians or an individual file for each custodian).

Real-World Use Case

A law firm used Clearwell for e-discovery 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. 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 messages between eight lawyers (that 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 dataset. 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 by discussion threads, and then by individual messages. 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 exported only the responsive email messages in native 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 e-discovery costs are spiraling out of control. Companies are under immense pressure to contain costs, more efficiently use internal resources, and meet court-appointed deadlines. Companies are reducing this pressure by adopting a new approach to e-discovery. By using Clearwell to process, analyze, search, and cull case datasets, companies are transforming their e-discovery processes. The end result: companies determine better case strategies, lower processing costs by up to 80%, reduce review workload by up to 90%, and regain control of their e-discovery process.

1. 2005/2006 Fulbright and Jaworski Survey available at http://www.fulbright.com/
2. Brian Ingram, "Locate Smoking Guns Electronically." COM.

Clearwell

© 2008 Clearwell Systems, Inc., Clearwell E-Discovery Platform is a trademark of Clearwell Systems, Inc. All rights reserved.