Clearwell

Best Practices for Streamlining Digital Investigations

A Clearwell White Paper

Table of Contents

Key Challenges Facing Digital Investigations Today 3

Limitations of the Traditional Investigations Process 4

Corporate Best Practices for Streamlining Digital Investigations 6

Summary 8

Key Challenges Facing Digital Investigations Today

Corporate investigators are increasingly challenged to perform more investigations in less time as they face greater demands from inside and outside the enterprise. Overall, businesses expect a 40% increase in the number of disputes they will face in the coming year.[1] From inside the enterprise, investigations related to HR issues, suspected fraud, and suspicious activity continue to strain resources. External forces such as changes in the regulatory environment, updated mandates for Sarbanes-Oxley, and the Department of Justice are fueling regulatory inquiries. For example, in 2009, there were 120 Foreign Corrupt Practices Act- (FCPA) related investigations, compared with 38 two years ago.[2]

With the increasingly pervasive use of email and electronic documents in all forms of business communication and activities, the amount of electronically stored information involved in an investigation is greater than ever before. From 2007 to 2011, businesses experienced a 10-fold increase in digital data growth, with 80 percent of that growth attributable to unstructured data.[3] Over 60 billion emails are sent daily throughout the world, and 90% of all documents generated today are electronic.[4]

Further, deadlines are increasingly aggressive and cost control remains a pressing concern. Investigators are under mounting pressure to rapidly find the key facts of an investigation so the appropriate course of action can be taken. While workloads are increasing, cost constraints are forcing investigators to make do with their existing resources, meaning they simply have to do more with less.

As a result of these challenges, investigators are taking a closer look at their existing workflows and searching for ways to immediately get to the heart of their investigations in a simple, efficient way. Realizing that their existing methods don't enable them to meet tomorrow's challenges, investigators are seeking new technologies that help streamline their workflows and enable them to rapidly convert mountains of data into actionable information critical to their investigations.

Limitations of the Traditional Investigations Process

Investigations conducted today using traditional technologies and workflows generally require a high degree of manual intervention, repetition, and brute-force, increasing the duration of an investigation and the risk of errors. Below is an example of a typical workflow applied to an investigation.

STEP 1: COLLECT DATA.

Documents from custodians involved in the investigation are collected and preserved as evidence. Since business records are contained within multiple data servers, companies must use a variety of standalone tools (i.e., Microsoft Outlook®, Guidance EnCase,® AccessData Forensic Toolkit®) to collect data from multiple data sources. These tools typically have arcane interfaces that are slow and cumbersome to use. In addition, container files usually need to be extracted to loose files and manually exported from one tool to another prior to processing, contributing to workflow inefficiencies and errors.

STEP 2: AUDIT AND FILTER COLLECTION.

Collected data is filtered by custodian, date range and file type. Key stages typically include the following:

- De-NIST data to remove non-user generated files irrelevant to the investigation
- Apply basic filters (i.e., date range and file type) to focus the scope of the investigation.

Data is generally not de-duplicated across the disparate data sources nor are files extracted from their container files until indexing. Consequently, investigators lack an immediate sense of the volume of data they are dealing with and how many resources they will need for the investigation.

STEP 3: SEARCH AND ANALYZE DATA.

Simple keyword searches and analyses are conducted using different tools that offer only limited search capabilities and have not been designed specifically for investigators. In general, investigators:

- Search for relevant data for each custodian within one application – and then must repeat the same search in each of the other applications. The inability to search and analyze data across all custodians at once forces investigators to repeat their tasks across data sets and analyze the same documents multiple times.
- Re-run searches time and again. This may be because of limited application functionality (i.e., previous searches couldn't be saved and used for subsequent searches) or because incremental new data has been added to the investigation, (i.e., naming of additional custodians or expansion of date ranges). Either way, the result is the creation of unnecessary workflow redundancies.
- Waste time sifting through thousands of mostly irrelevant documents returned by searches, a drawback of the traditional tools' inability to help investigators bulk eliminate false positives.

- Answer the frequently asked question of who said what to whom only by manually and painstakingly re-creating the flow of events, using complex diagrams depicting relationships with time/date stamps of emails connecting senders and recipients. Other custodian-document relationships are also separately manually analyzed to understand the timing and flow of communications across different custodians. The lack of visual analytics that help uncover event chronologies and situational contexts limits an investigator's ability to quickly identify and assess case facts.

In addition, investigators are unable to perform advanced analyses based on:

- A universal view of all variations of a custodian's email address with the choice to designate specific addresses upon which to focus
- Inbound or outbound email communications for specific custodians to understand the flow of key information
- Files that have been saved under a different name or file type, since they can't be identified or traced back to their original custodian.

As a result, investigators face high false positive and low cull-down rates that necessitate additional searches and reduce data by only 30% to 40%. More importantly, investigators risk an inaccurate and incomplete investigation because they can't easily find all the relevant evidence or the smoking gun.

STEP 4: DELIVER DATA.

With traditional tools, investigators must spend an inordinate amount of time manipulating files from multiple point products to develop a standardized format for data delivery to case requestors or other members of the investigation team (i.e., paralegals, inside counsel or HR) for further review. Oftentimes, gigabytes of analyzed documents from each application are printed out or separately saved to DVDs and physically delivered to business team members for review. The lack of a single, Web-based, self-service portal for the team to quickly review relevant documents results in unnecessary overhead and complexity.

These traditional workflows severely limit an investigator's ability to quickly cull out irrelevant data, intuitively unearth critical documents, and administer investigations with ease and simplicity. Progressive investigators need a new approach that can help them streamline their investigations with greater accuracy, speed, and effortlessness than ever before.

Corporate Best Practices for Streamlining Digital Investigations

The Clearwell E-Discovery Platform is the first enterprise-class investigation management solution designed and proven to accelerate investigations within a single, easy-to-use application. Drawing upon corporate best practices of hundreds of enterprises, Clearwell provides an integrated processing, search, analysis, review, and delivery platform. With Clearwell, investigators are up and running within 25 minutes and can rapidly cull-down data and quickly expose and prioritize key facts of an investigation in context through advanced analyses. Clearwell's Web-based portal also allows investigators to easily administer and manage access to their investigations, so they can more effortlessly scale and increase productivity.

	Traditional Investigations		Clearwell Digital Investigations	
COLLECT DATA	Export data manually across tools	1-2 Days	Extract data automatically from logical files	2 Hours
AUDIT and FILTER	Guess at data volumes and resource requirements arbitrarily	1 Days	Estimate budgets and timelines systematically and accurately	1 Hours
SEARCH and ANALYZE	Spend costly cycles in search of relevant facts	1-2 Days	Gain immediate visibility into case facts	5 Hours
DELIVER DATA	Prepare data manually for review	2-5 Days	Provide instant Web-based access to case data	0 Hours
	TOTAL TIME: 5 to 10 DAYS		TOTAL TIME: 1 DAY	

The Clearwell E-Discovery Platform draws upon corporate best practices to accelerate digital investigations.

Revisiting the earlier example of the step-by-step investigation workflow, the following process details how real companies are employing the Clearwell E-Discovery Platform to streamline their digital investigations.

STEP 1: COLLECT DATA.

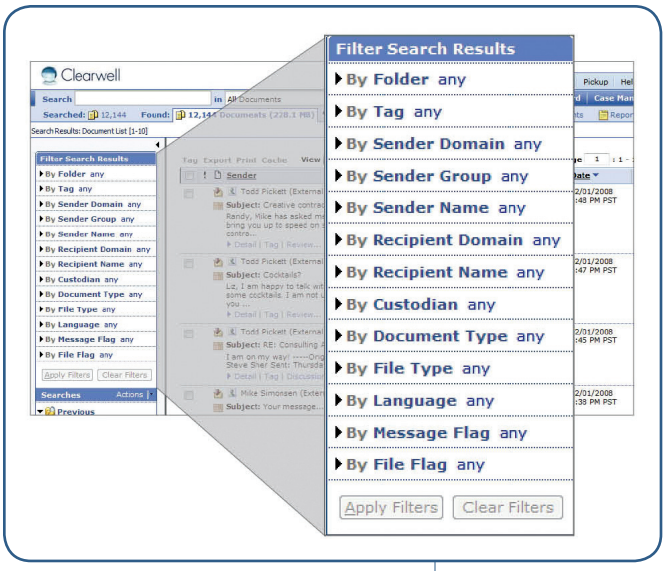
This step remains largely unaltered since data must still be collected from multiple data sources. However, once this is completed, Clearwell seamlessly integrates into a company's current collection process. Since Clearwell can automatically extract data from logical files, manually exporting the data from the standalone tools to Clearwell is unnecessary and the collection process is streamlined.

Clearwell Pre-Processing Filters can reduce case datasets by up to 30% by filtering on custodian, data, strong file type, and file size prior to processing.

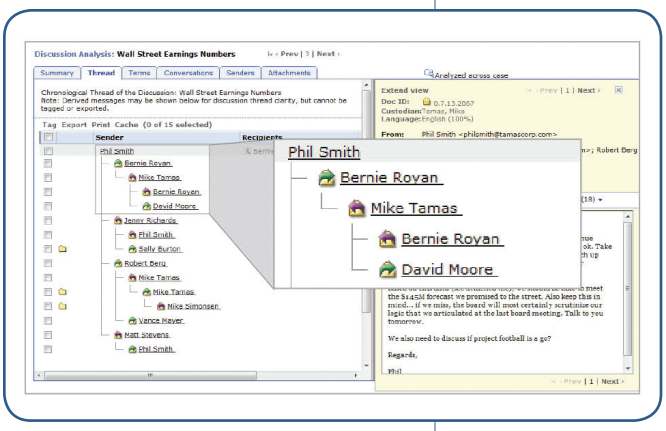
STEP 2: AUDIT AND FILTER COLLECTION.

Instead of jumping directly into data indexing, Clearwell's Pre-Processing Module offers a set of interactive reports, which provide investigators with detailed visibility into the makeup of the data collection within a case. Investigators can then audit the collection for missing data and accurately filter out irrelevant files prior to full indexing, significantly reducing data volumes. This allows users to:

- Utilize visual analytics to summarize overall document set characteristics and present detailed analysis by custodian, timeline, and file type. This rapidly confirms that all case data has been collected and allows for accurate estimation of investigation budgets and timelines, as well as an increasingly defensible process.
- Interactively filter collected data by custodian, date, strong file type, and file size prior to full processing. Clearwell also provides one-click filtering of custom file and "NIST List" items from Guidance LEF and E01 containers, eliminating the need to manage and manually export data to another tool to perform this task



Clearwell Auto-Filters automatically group search results by metadata fields such as tag, sender domain, recipient domain, document type, custodian, and language type.

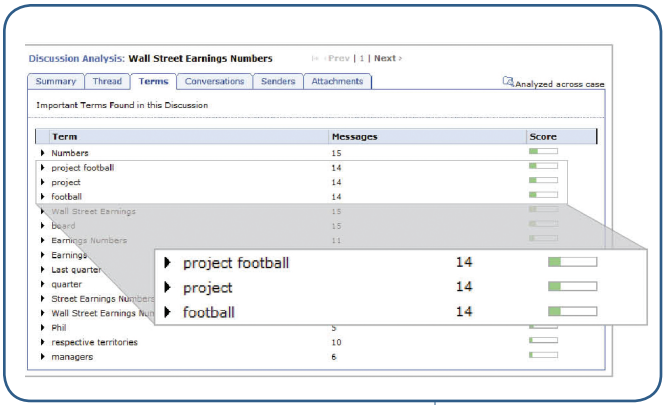


Clearwell Discussion Threads dynamically link together all related messages into chronological threads that capture entire discussions to determine exactly who knew what and when.

STEP 3: SEARCH AND ANALYSIS.

Once full indexing is complete, Clearwell's advanced search and analysis capabilities enable investigators to gain immediate visibility into case facts.

- **Interactively reduce collected data.** Clearwell's Auto-Filters automatically group the data set by several metadata fields such as email sender domain, sender name, email recipient domain, recipient group, recipient name, document type, and language type, and display exact hit counts across the entire search result set for every filter. This allows investigators to reduce data by 80% to 90% by simply clicking on a checkbox to cull-down documents containing irrelevant information - for instance, removing all extraneous emails received from domains like amazon.com or espn.com.
- **Route special documents to experts.** As one example, Clearwell automatically identifies documents written in different languages, which can allow an investigator to quickly route Chinese email and documents to a translator or to another investigator fluent in Chinese.
- **Discover exactly who knew what and when.** Clearwell's Discussion Threads analyze both metadata and content to link together emails, including all replies, carbon copies, blind carbon copies and forwards, into chronological threads. By tracing the thread, users can quickly identify all primary and peripheral participants involved and address the most important question in an investigation: who knew what, when.
- **Improve the accuracy of searches.** Clearwell's Participant Selector provides a list of email addresses associated with a custodian, allowing investigators to specifically choose the addresses that should be included in the investigation.
- **Analyze individual communication patterns.** Clearwell's People Analytics enable investigators to analyze individual and group-to-group communications within a company or with customers, suppliers, and partners. Users can easily access a ranked list of top custodians for a search or monitor communications between regulated and non-regulated divisions.
- **Discover secret project names and code words.** Clearwell's Term Analytics analyze noun phrases to help investigators find code words or secret project names that are likely to be relevant to the case.
- **Find all instances of a document.** Clearwell's File Analysis capabilities automatically find all instances of an attachment or loose file based on multiple hash values that relax the last modified date, strong file name, or other metadata. Investigators can then find all instances of a document even if a document name or extension has been changed.



Clearwell Term Analytics analyze noun phrases to uncover secret project names and code words.

STEP 4: DELIVER DATA.

Whereas traditional tools often require manual and time-consuming processes to deliver data, Clearwell enables investigators to:

- **Provide access to case data through a single Web-based portal.** Clearwell's Web-based platform provides a single, easy-to-use portal that allows investigators, business users, paralegals, counsel, law enforcement, and other authorized users to access case data in real time.
- **Enable self-service for reviewers.** Clearwell provides reviewers with the ability to quickly access key facts in standard, easy-to-read formats directly.

Using traditional technologies, the average time required to collect and cull-down investigation data to its relevant facts and deliver the data to reviewers can take between five and ten days. By contrast, using Clearwell investigators can generally complete this process in just one day, dramatically accelerating the rate at which investigations can be completed

Summary

As today's corporate investigators face increasing pressure to complete more cases involving growing amounts of digital evidence with less time and money, they can no longer be satisfied – or effective – with traditional, incomplete, and inefficient tools. In addition to prolonging an investigation, these traditional approaches introduce greater risk and costs that investigators cannot afford. Rather, investigators need a next generation digital investigation solution that enables them to gain immediate visibility into case facts easily, simply, and intuitively. The Clearwell E-Discovery Platform is uniquely designed and built to meet the needs of investigators and is being used by hundreds of companies across thousands of cases to improve their investigative process. With Clearwell, investigators can tackle their cases with unparalleled ease and confidence.

Learn more about how the Clearwell E-Discovery Platform is being used to accelerate investigations by viewing Clearwell's On-Demand Webinars:

- Winning Strategies for Successful Digital Investigations, and
- The Analysis-Powered Investigation: How the New Clearwell E-Discovery Platform 5.0 Can Find the Smoking Gun

Both can be found in the Resource Library at www.clearwellsystems.com.

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

1. Fulbright & Jaworski, "6th Annual Litigation Trends Survey Report," 2009
2. Gibson, Dunn & Crutcher LLP, "2009 Mid-Year FCPA Update," July 2009
3. Channel Web, "XChange 2010: IBM Counts On Software As Key Digital Future," March 2010
4. Sally Kane, "E-Discovery Explosion: E-Discovery Growth and Challenges"

Clearwell