

# Clearwell Processing & Analysis Module

The Clearwell Processing & Analysis Module provides the ability to rapidly and accurately filter, process, search, and analyze data in multiple formats and languages. Using Clearwell, corporations, government agencies, and law firms perform early case assessments and rapidly cull down data, thereby reducing overall e-discovery costs. As an integrated part of the Clearwell E-Discovery Platform, the Processing & Analysis Module supports the iterative workflows required during real-world e-discovery, delivering deep insight into case facts and enabling a new level of transparency and defensibility throughout the e-discovery process. Key features of the Processing & Analysis Module include:

## 1. Processing

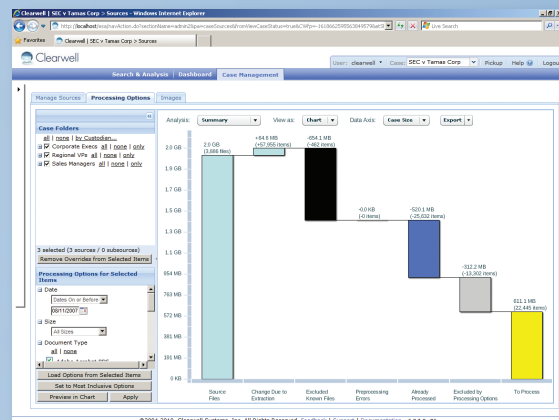
**Advanced Pre-Processing Filters:** Enable users to interactively filter data by custodian, date, strong file type, and file size prior to processing. Clearwell also provides one-click filtering of custom file and "NIST list" items, significantly reducing downstream processing and review costs.

**Pre-Processing Analytics:** Visually summarize overall document set characteristics and presents detailed analysis by custodian, timeline, and file type. This rapidly confirms that all case data has been collected and allows for accurate estimation of e-discovery budgets and timelines.

**Rapid Processing:** Utilizes a high-performance architecture to process and intelligently deduplicate documents at rates of up to 1 TB per day.

**Robust File Support:** Processes and analyzes over 400 different file types from data stores and network file shares including Microsoft Office® and PDF documents, various email formats such as PST, NSF, MBOX, OST, and EMLX, attachments, and Guidance Software® LEF and EO1 files. Translates text in image files to searchable content with integrated OCR and identifies over 40 types of hidden content.

**Multi-Language Support:** Provides full Unicode compliance and supports English and 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.



**Pre-Processing Analytics:** Graphically depict data volume, file types, and time frames of collected data prior to processing.

## 2. Analysis

**Discussion Threads:** Utilize patented algorithms to dynamically link together all related messages into chronological threads that capture entire discussions, including all replies, carbon copies, blind carbon copies, and forwards. This enables users to quickly identify all participants and determine who knew what and when.

**Find Similar:** Allows reviewers to easily identify and view emails, attachments, and loose files with similar content to the document under review. Reviewing similar documents together accelerates the review process and ensures greater tag consistency.

**Topic Classification:** Utilizes patent-pending linguistic algorithms to automatically organize documents into specific topics, enabling users to quickly analyze all documents related to a particular subject.

**People Analytics:** Analyze individual and group-to-group communications within a company or to customers, suppliers, and partners. Users can easily access a list of top custodians for a search or monitor communications between divisions.

**Term Analytics:** Leverage natural language algorithms to analyze noun phrases, which help users uncover secret project names and code words that may be relevant to a case or investigation.

### 3. Search

**Transparent Search:** Delivers a set of search features that enables a more collaborative and defensible e-discovery search process.

**A. Search Preview:** Provides visibility into matching keyword variations prior to running a search. Users can selectively include relevant variations or exclude false positive variations from their search query, thereby adding relevant documents and removing irrelevant documents from search results.

**B. Search Filters:** Enable real-time search result filtering for individual queries or variations and allow users to sample the filtered documents to evaluate the effectiveness of their search as recommended by The Sedona Conference®.

**C. Search Report:** 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.

**Multi-Keyword Search:** Delivers the ability to run up to 100 queries simultaneously, dramatically decreasing the time needed to evaluate the effectiveness of keyword searches.

**Advanced Search:** Provides the ability to construct advanced searches based on numerous metadata and derived fields. It supports both stemmed and unstemmed (literal) searches and provides power-user capabilities including Boolean, wildcard, fuzzy, and nested proximity searches.

Variation	Matching Emails	Matching Unique Files
<input type="checkbox"/> falsestart	274	4
<input type="checkbox"/> falseetto	74	5
<input type="checkbox"/> falsie	0	5
<input type="checkbox"/> falsification	1	6
<input type="checkbox"/> falsifications	15	4
<input checked="" type="checkbox"/> falsified	68	8
<input checked="" type="checkbox"/> falsify	27	0
<input checked="" type="checkbox"/> falsifying	29	8
<input type="checkbox"/> falsity	148	5
<input type="checkbox"/> falsom	43	2
<input type="checkbox"/> falstaff	70	2
<input type="checkbox"/> falstore	25	4

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

**Auto-Filters:** Automatically group search results by metadata fields such as tag, sender domain, document type, custodian, and language type and display exact hit counts across the entire search result set for every filter.

### 4. First-Pass Review & Export

**Optimized Review Interface:** Delivers an HTML review interface that maximizes screen real estate and minimizes mouse clicks to increase review throughput. Dual monitor support, breadcrumb navigation, and thumbnail view allow reviewers to quickly navigate through document sets.

**Decision-Tree Tagging:** 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.

**Load File Configurator:** Provides customizable load file creation during export. Expanded options allow for multiple metadata formats CSV, DAT, EDRM XML, as well as Concordance Relativity®, and Summation® load files.



Clearwell Systems, Inc.  
441 Logue Avenue  
Mountain View, CA 94043  
650.526.0600 tel  
650.526.0699 fax  
www.clearwellsystems.com

To learn how Clearwell can deliver value to your organization, contact an e-discovery expert at: [info@clearwellsystems.com](mailto:info@clearwellsystems.com), or call us at: **877.727.9909**.