Veritas eDiscovery Platform™

Case Administration Guide

10.1.1

VERITAS™

Veritas eDiscovery Platform™: Case Administration Guide

The software described in this book is furnished under a license agreement and may be used only in accordance with the terms of the agreement.

Last updated: 2022-2-9.

Legal Notice

Veritas Technologies LLC
2625 Augustine Drive
Santa Clara, CA 95054
http://www.veritas.com

:

Contents

:

:

:

Case Administration Guide

The Case Administration Guide provides administrators of the product with details on how to set up and manage cases. It also describes performing pre-processing through post-processing tasks in preparation for end users to search, review, and analyze the data. This guide also provides details on the handling of various file types and hidden content.

This guide is intended for users with the **Case Admin** role, the **Group Admin** role, or the **System Manager** role.

This section contains the following:

About This Guide

Case Administration Guide is intended for **case administrators**, decision makers, and anyone who is interested in understanding how to prepare and process data in a case through the various stages of the Veritas eDiscovery Platform. For information about administering the system, refer to the System Administration Guide.

Revision History

The following table lists the information that has been revised or added since the initial release of this document. The table also lists the revision date for these changes.

Revision Date	New Information
February 2022	• Updated version for release 10.1.1 • Updated the information on allowed characters in redaction set name while creating a redaction set. See *"To set up redaction sets" on page 139*.
December 2021	• Updated version for release 10.1 • Updated the image of Settings screen in *"Defining New Cases" on page 18* to include Persistent Hit Highlighting feature. • Added a note about Native viewer performance. See *"Persistent Hit Highlighting" on page 25* • Added a note under *"Selecting a Case" on page 177* to refer to the "Case Administration Workflow Recommendations" if a warning message to update checksum for emails is seen on selecting a case.
July 2021	• Added information on the Persistent Hit Highlighting feature

Revision Date	New Information
March 2021	• Updated information related to redactions, production, and exports • Updated the Production Folder Settings section • Minor edits
March 2020	• Minor edits
October 2018	• Added information on imaging of hidden content in Microsoft Excel, Word, and PowerPoint. See *"Hidden Content" on page 241*. • Minor edits throughout
March 2018	• Configurable option to generate a slip sheet for excluded items • Processing of partially converted OST files
December 2017	• Auto-conversion of OST to PST data files • Annotation • Bulk Redaction • Information Classification (VIC) • Preset Reason Code Tags
June 2017	• AD sync enable/disable • When to run Update Checksum for Email tool • Update community link • Minor edits throughout
July 2016	• Added information on Access Groups feature. • Added information on new user roles "**System Manager**" and "**Group Admin**". Explained how **Case Admin** role has changed. • Changed references to converting OST mail formats to "supported formats". • Branding and minor edits.
August 2015	• Added location option to Exceptions and Imaging and Rendering exports. • Changed Processing > Access Groups tab to Processing > Groups. • Added support for EML format of Journal messages. • Remove Rights Management Guide.
March 2015	• Image accessibility. • Added Production Membership Report. • Office 2013 support. • Branding and minor edits.

Revision Date	New Information
October 2014	• New case setting for number of attachments to show per message. • Case Automation. • Updates to Review Dashboard. • Added OCR languages. • Added System Image tags. • New Slip sheets. • Update System > Users > Groups tab to Access Groups. • Deprecated Predominant Language feature. • Added guidelines for text block exclusions during Processing/Search. • Added changes for "Separately produce embeddings" checkbox. • Branding and minor edits.
March 2014	• Added RMS enable/disable case support and RMS superuser ID.
December 2013	• New Audio Search Default Language case settings. • Added EV10.0.4 note regarding Sharepoint 2013 and Exchange 2013. • Added new docs to preface.
June 2013	• Image Remediation - bulk image end-to-end processing (replaces "Importing TIFF Image Files). • Image Overlay - individual image overlay in review mode. • New Reports UI tab for Pre-Processing and Processing reports. – New and revised reports. • Ability to specify, at the case level, contacts, calendar items, tasks, journal entries and posts (files). • Customize slip sheet printed text for productions. • Container handling enhancements. • Mixed mode production results analysis. • Efficiency improvements for PST/NSF crawling. • General updates to graphics, text and minor edits. • New columns for home appliance case selection. • Processing setting "Extract documents from container files" replaced. • Added case matter fields (Case Caption, Key Dates, etc.) to new cases. • Added container file IDs. • Separate tagging of attachments enhancements (item and document family). • Enterprise Vault journal envelope message settings. • Email Header View and Search settings.
Sept 2012	• Attachment-level tagging option added to *Defining Tag Sets* procedure. • Native Imaging and large document imaging analysis for Production folder. • Release template updates and edits.

Revision Date	New Information
March 2012	• Branding and edits. • This book now contains all the content from the following documentation sources: – *Multiple Language Support* – *Pre-Processing Navigation Guide* – *Processing Exceptions Reference* – *Supported File Types* • "Discovering Archive Sources" section moved to *System Administration Guide*.
Feb 2012	• New Monthly Billing Model configuration / license usage option.
Nov 2011	• New top menu navigation, case selection, and **Case Admin** workflow. • Load file import and load file import pre-processing report. • Integrated Review Dashboard to view all activity for a case.
May 2011	• Concept Search configuration. • Processing and pre--processing enhancements: – batch-level reporting for discovered data. – option to disable "Hide Informational" warning messages (enabled by default), or change property default setting. • Moved Advanced Export information to Export and Production Guide.
Feb 2011	• Scalable Folder Management and user interface enhancements. • Additional security and administrative options. • User deletion. • Resubmit documents for OCR.
Dec 2010	• Documented new features: – custodian merge/unmerge. – Find similar (noted user-configurable threshold). • Added graphics and description for pre-processing feature enhancement - viewing errored files during pre-processing. • Inserted description for new case setting options: – Document Duplication in Milliseconds. – Process Truncated Lotus Notes Documents. • (Minor revisions and graphics updates throughout).

Technical Support

Technical Support maintains support centers globally. All support services will be delivered in accordance with your support agreement and the then-current enterprise technical support policies.

For information about our support offerings and how to contact Technical Support, visit our website:

https://www.veritas.com/support

You can manage your Veritas account information at the following URL:

https://my.veritas.com

If you have questions regarding an existing support agreement, please email the support agreement administration team for your region as follows:

Worldwide (except Japan)	CustomerCare@veritas.com
Japan	CustomerCare_Japan@veritas.com

Documentation

Make sure that you have the current version of the documentation. The latest documentation is available from:

- **Documentation** link at the bottom of any page in the Veritas eDiscovery Platform landing page.
- **Veritas Products Web site:** https://www.veritas.com/product/a-to-z

Documentation Feedback

Your feedback is important to us. Suggest improvements or report errors or omissions to the documentation. Include the document title, document version, chapter title, and section title of the text on which you are reporting. Send feedback to:

eDiscovery.InfoDev@veritas.com

You can also see documentation information or ask a question on the Veritas community site.

https://vox.veritas.com/

Preparing Your Case

For information about how to create new case, refer to the following topics:

About Access Groups and Roles

Starting with Veritas eDiscovery platform release 8.2, the optional Access Groups feature provides a significant level of access control. Case access can be authorized individually, or by Access Groups, across the entire workflow.

An Access Group is a group of users who generally need to share files, workflows, and cases. Access Groups can be used to *prevent* functional units from sharing cases, where appropriate. If you have operating units or entities that are not involved in the same cases, searches, etc., you may wish to assign them to different Access Groups.

If your organization has different departments that are typically involved in the same cases, they should be in the same Access Group. However, you do not need to use Access Groups, or assign any cases or users to an Access Group.

Assign Access Levels to Groups

You can restrict or assign access levels to the following group entities in the eDiscovery platform:

- Users
- Cases
- Legal Holds
- Sources
- Locations
- Collection Sets

Groups Considerations When Creating New Users

- When creating new users, you must choose Access Type when the user is created. You can assign them to an access **group**, or to authorize them only to specific **cases**.

  Note: You cannot give a user both group access and case authorization. For users with both in version 8.1.1, after an upgrade to 8.2, existing users will retain their case authorization but lose their group assignments. Changes that happen as a part of the upgrade process are reported in the upgrade logs. See the *Upgrade Guide* for more information on logging and reporting.

- Starting with version 8.2, only users with the **System Manager** role, the **Group Admin** role, or the **Case Admi**n role can add users, assign them to roles, assign them to Access Groups, or assign them to cases. A **System Manager** can assign a user to any role, and to any access group. A **Group Admin** can only assign users to their group. A **Case Admin** can only assign users to their case.

- When you create a new user, they will have all of your group access rights by default.

  Note: If a **System Manager** creates a new user and does not authorize either specific cases or place the user in specific Access Groups, that new user has access to all cases.

- When creating new users, you can choose to place them in an access group, or to authorize them only for specific cases. If you do not do one or the other, the new user will have access to either the same access groups as you, or the same cases.

  Note: In previous versions, when creating a new user the default was no visibility into cases until an explicit assignment was made.

- If you wish to create cases but keep them invisible temporarily, you can create an Access Group with few or no users, and place the case there. It can be moved later.

- When users are granted access on a case-by-case basis, this is called **Case-Authorized** operation. Users and cases function as described in *"Case Authorized Operation" on page 15*.

Note: If you are upgrading and are currently using Access Groups, you and other stake holders should decide how to arrange user and case authorization to Access Groups before upgrading. At least one person must have the **System Manager** role to work with Access Groups. As a system administrator, you should work with the **System Manager**, any other **Case Admins** and

the workflow management team to define Access Groups, role assignments, and how to resolve related issues as part of any upgrade. Refer to the *Veritas Upgrade Overview* for more information.

Case Authorized Operation

If you choose not to assign users to an Access Group, these users will be in a common pool and will have access to all cases. Users can then be authorized for specific cases. This is called **Case Authorized**. Restrictions on case access can then be further restricted on a user-by-user basis. Case Authorized operation is a less flexible alternative to Group Access. It is recommended for when a user should have access only to a small number of cases.

Note: When users are created, they are either assigned to all Access Groups (the default) or, if the Case radio button is selected, authorized for all cases, unless you specifically limit either Access Groups assignment or case authorization.

Case Administration Overview

With the addition of Access Groups, the functions of the previous Case Administration role are controlled through three roles: **System Manager**, **Group Admin**, and **Case Admin**. **Group Admins** have the same privileges as **Case Admins**, but in addition have certain Collection and Data Set privileges, and Legal Hold management privileges. The primary difference between **System Manager** and **Group Admin** is that **Group Admins** cannot create new groups. They see and control only the groups they were added to by the **System Manager**.

Roles

The available, pre-defined roles for users are:

- **System Manager**. Allows access to all system and case administration, search, and reporting functions. This gives unrestricted rights to manage the entire system, including admin-level access to all groups and cases.
- **Group Admin**. This gives the user the ability to add and remove users, cases, and other items from the group to which they have access, and to perform other administrative tasks.
- **Case Admin**. Allows access to all case administration, search, tagging, export, and reporting functions, but not System Manager functions.
- **Case Manager**. Allows access to one or more cases. It includes case admin rights, except for source setup, plus all case user rights.
- **Case User**. Allows access to most case search, tagging, export, and reporting functions, but provides no system or case administration functions.
- **eDiscovery Admin**. Allows a user to manage the Identification Data Map, perform Collections, and Process, Analyze, and Review. System Manager privileges are not included.
- **Collection Admin**. Allows a user to manage only the Identification Data Map, and to perform Collections.
- **Legal Hold Admin**. Allows user administrative access and management of Legal Holds.

Starting with version 8.2, only users with the **System Manager**, **Group Admin**, or **Case Admin** role can add users, give them roles, assign them to Access Groups, or authorize them for specific Cases. A user added by the **System Manager** can be given any role, and assigned to as many Access Groups as desired. A user added by the **Group Admin** can only be assigned to groups previously assigned to that **Group Admin**. A user added by the **Case Admin** belongs to that case only. For complete details on roles and their rights, refer to the *System Administration Guide: Defining User Roles*.

Roles and Rights

Overview of Case Rights (Refer to the *System Administration Guide* for a detailed description of these and other included rights.)	**System Manager Role**	**Group Admin Role**	**eDiscovery Admin**	**Case Admin Role**
Allow system management and support access. Allow Group creation and management. Allow admin user and role management.	Allowed	Not allowed	Not allowed	Not allowed
Allow user management, role definition.	Allowed	Within group	Not allowed	Not allowed
Allow access to integrated analytics, tags, dashboard, charts, and reports.	Allowed	Within group	Within group	Within case
Allow case setup, case status access, source setup, user management, tag definition, custodian and participant.	Allowed	Within group	Within group	Within case
Allow Case Home and All Cases Dashboard Access. Allow new case creation, case backup, restore, deletion, template creation.	Allowed	Within group	Within group	Not allowed
Allow collections access and management, data map management, backup and restore.	Allowed	Within group	Within group	Not allowed
Allow viewing, tagging, bulk tagging, export.	Allowed	Within group	Within group	Within case
Allow Legal Hold access and management.	Allowed	Within group	Within group	Not allowed

Note: The rights of a system-defined role cannot be altered. Users with the **System Manager** role can create a new role and modify its rights as needed.

Depending on your organizational and security requirements, you may need to change the roles of some **Case Admins** to be **Group Admins**, or **System Managers**. Work with an administrator who has **System Manager** role for initial group setup and role re-assignment across groups. Only the **System Manager** role can create Access Groups and give users the **Group Admin** role, or assign them to more than one group.

About the **Case Admin** Role

The **Case Admin** role is a case-level processing role, allowing access to one or more cases. The case must be within the Access Group, if groups are in use, or granted by **Case Authorization**.

General rights include (by default) access to integrated analytics, analysis tags dashboard, management charts, and reports. The **Case Admin** role also includes document access rights, that is, all **Analysis and Review** module activities.

Note: User access rights are defined by a combination of the role you give to the user, whether the user is in an Access Group, and any further authorizations or restrictions on case access that you make.

Applying Access Profiles for a User

Access Profiles can be used to further restrict which documents, folders, and tags within a case can be seen by which users. Permissions set by the user's role normally extend to all cases to which that user has access. **Case Admins** can further control case access by applying Access Profiles for each user authorized to access that case. A **Case Admin** cannot apply an Access Profile other than **Case Admin**, **Case Manager**, **Case User**, or a custom role. In other words, a **Case Admin** can further restrict access rights, but cannot increase them.

Access Profiles are applied on a case-by-case basis. For more information on Access Profiles, refer to *System Administration Guide: Defining Case Access Profiles*.

For more information about where to find functions specific to administrative user roles within the new user interface, refer to the *Veritas eDiscovery Navigation Reference Card*.

Case Administration Workflow Recommendations

Avoid Simultaneously Running Case and Update Checksum for Emails Jobs

Make sure that no Update Checksum for Emails jobs are running before attempting to run any jobs (for example discovery, processing on a case folder, review or export) for the case.

The Update Checksum for Email job runs within the context of a case and if it runs concurrently with any case jobs, there may be issues.

To check if the update checksum job is running

1. From the navigation bar, go to **System > Jobs** and select **All Jobs** under the Context.
2. Look at the **Status** column for "**Update email checksums**" jobs.

Note: For more information, see *"Update Checksum for Emails" in the System Administration Guide* and the technical article: "How to resolve Checksum Duplication issues using the 'Update checksum for emails" located:

https://www.veritas.com/content/support/en_US/article.100051881.

Defining New Cases

To get started working with a set of documents in the application, you need to create a case. You must have the **Group Admin** or **System Manager** role to create a new case.

Case Workflow

The product enables a fully integrated end-to-end case workflow which allows users a single overall view of one individual case at a time, and all activities pertaining to that case. For example, a case may have started with a legal hold (Legal Hold module), then collection data was added (Identification and Collection module), before being processed (Processing module), and prepared with batches for review (**Analysis and Review** module). See also *"Using the Review Dashboard" in the User Guide*. At any point, administrators (with appropriate access) can view the overall status of the any case at each of these steps during the case workflow process.

For more detailed information about each of these functions and modules, refer to the following guides:

- **Legal Hold**: Refer to the *Legal Hold User Guide*.
- **Identification and Collection**: Refer to the *Identification and Collection Guide*.
- **Additional Pre-Processing and Processing Detail**: See the sections *"Pre-Process Your Source Data" on page 67*, *"Pre-Processing Navigation" on page 145*, and *"Processing Exceptions" on page 159* in this guide.
- **Analysis and Review**: Refer to the *Veritas eDiscovery Platform User's Guide*
- **Audio Search:** (For customers licensed for audio content): Refer to the *Audio Search Guide*.

Note: Before you attempt case setup, make sure you have a licensed and installed Pre-Processing module with pre-processing enabled on your system. These prerequisites are necessary to later analyze your pre-processed data, and view advanced pre-processing options and filters. If you fail to enable pre-processing at case startup, you will not be able to process LEF files, de-NIST loose files, or get Sent dates in email files (PST, MSG, EML, and NSF). For more information, refer to the section *"Pre-Processing Navigation" on page 145*.

After you create a case, you can define the sources of the documents that you want to index and analyze, as well as other case-specific features, such as folders, tag categories, and participant access groups.

Before You Begin: You must have the **System Manager** role or the **Group Admin** role to create cases and case templates. The **Case Admin** role can define document sources and case settings. Alternatively, as a **Case Admin**, you can simply apply an existing case template, if any, to specify case settings.

Note: If you later make changes to a case template, the case(s) to which that template had previously been applied do not update. Their case settings must be changed separately, if necessary.

To create a new case

1. From the navigation bar, click **All Cases > New Case**. (Alternatively, from the drop-down menu, select **Create a new case**.)
2. Specify the following information. An asterisk (*) indicates a required field.

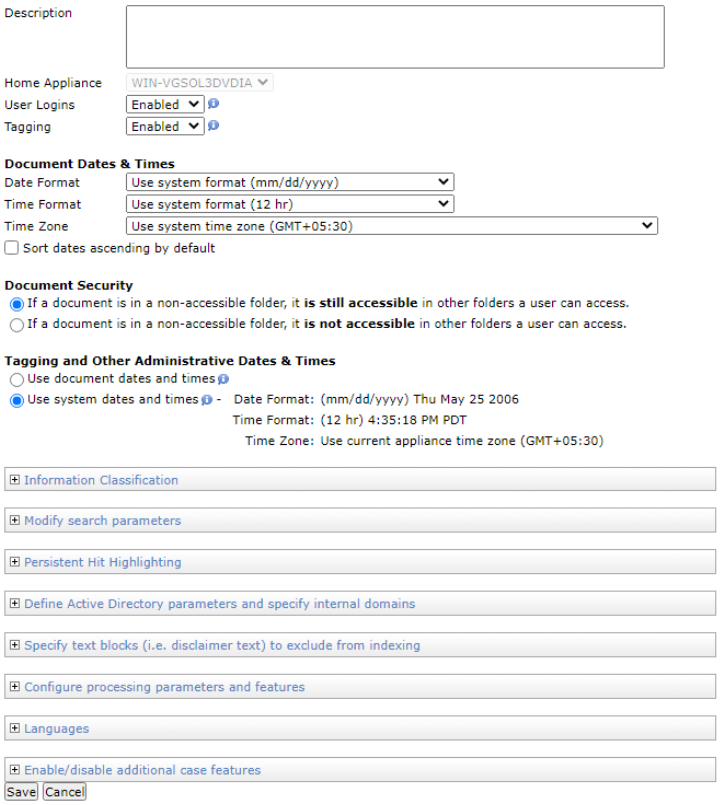
New Case Information

Field	Description
General	
Name*	Enter a case name (up to 35 characters).
Description	Enter a description of the case (up to 255 characters).
Access Groups	If you are using Access Groups, select the group(s) to which the case should belong. Note that by default the case is in all groups. If no group is specified, the case will be seen by all users, with access type **Group**, but not by users who have been authorized specific cases only (**Case Authorized**).
Number	
Type*	Select the type of case from the drop-down menu. **Note:** Users with the **System Manager, Group Admin, eDiscovery**, **Collection**, and **Legal Hold Admin** role can add or edit case types in the **All Cases > Settings** screen.
Business Unit	Enter the company's business unit or name to be associated with this case. **Note:** Users with the **System Manager, Group Admin, eDiscovery**, **Collection**, and **Legal Hold Admin** role can add or edit case types in the **All Cases > Settings** screen.
Enter more case information...	
Case Caption	Enter a caption for the case (up to 255 characters).
Status	Enter the status of the case from the drop-down menu. **Note:** Note: Users with the **System Manager, Group Admin, eDiscovery**, **Collection**, and **Legal Hold Admin** role can add or edit case types in the **All Cases > Settings** screen.
Court/Jurisdiction	Enter the Court/Jurisdiction of the case from the drop-down menu. **Note:** Note: Users with the **System Manager, Group Admin, eDiscovery**, **Collection**, and **Legal Hold Admin** role can add or edit case types in the **All Cases > Settings** screen.
Docket Number	Enter a docket number for the case (up to 255 characters).
Key Dates	
Filed Served, Court Close Date	Enter the dates around the case to identify key dates in the case. **Notes:** Cases created in versions prior to 7.1.2 Fix Pack 2 will have "Start Date" mapped to "Served", and "End Date" to "Close Date"**.**

New Case Information (Continued)

Field	Description
Staffing (All names entered will be considered Team Members)	
In-house Counsel Outside Counsel Lead Attorney Lead Paralegal	Enter the names of users or team members to be as these special types of team members in this case. The product automatically adds the names to the list, and can be re-used for future cases. **Note:** The names entered will be considered team members.
Other	
Case Notes	Enter any case notes for the case (up to 255 characters).
Custom Case Fields	Enter any custom case fields for the case. (up to 255 characters). **Note:** There are an unlimited number of custom case fields that can be added to a case.
Setup	
Home Appliance	If you have a cluster of appliances; select the appliance where the case is stored. To determine the best home appliance for the case, consult the information displayed in the columns: **Free Disk Space**, **Cases,** and **Indexed Docs**. Best practice is to assign the case to the appliance with the most free disk space. The appliances will be sorted so that those with the most free disk space will be at the top of the list.
Process Settings Template	If you have defined one or more templates to be used for this case, select the appropriate template from the drop-down menu. **Note:** If you have the **System Manager** or **Group Admin** role, you can create templates for cases, using the steps in *"Defining Case Templates" on page 198*.

3. Click **Save and Set Up Processing**. The **Processing > Settings** screen displays the new case.



4. Specify the following case settings. An asterisk (*) indicates a required field.

5. Click **Save** to submit the new case, or click **Cancel** to discard your changes.

Next Steps:

If you are using Access Groups, double-check that the case is in the correct Access Group.

To specify the document sources for the case, see *"Selecting Document Sources and Pre-Processing" on page 49*

New Case: Processing Settings

Field	Description
Description	Enter (or re-enter) a description of the case (up to 255 characters), even if you already entered one on the previous screen.
Home Appliance	(Once selected, the appliance cannot be changed.)
User Logins	Select **Disabled** to prevent non-administrative users from accessing the case. You can enable user access after the initial configuration and indexing are complete.
Tagging	Select **Disabled** to prevent all users from tagging documents in the case.

New Case: Processing Settings (Continued)

Field	Description
Document Dates and Times	Document-specific date/time settings are useful when the documents in a case originate in a different time zone from the location of the appliance. Each case can have its own document date and time settings, thereby allowing a single appliance to support cases originating from multiple locations. For example, a law firm headquartered in New York, which has its system-level date and time settings set to a US date format and Eastern time, may be managing a case with documents that originated in London. The system time zone is U.S. Eastern time and the format is based on the 12-hour clock. To enable reviewers to see document dates and times as the London custodian would see them, the administrator configures the following document settings: • **Date format**—dd/mm/yyyy • **Time Format**—24 hour • **Time Zone**—GMT With these settings, all document-specific information in the case is displayed in the document (London-GMT) time zone using the 24-hour clock. In addition, the European date format (dd/mm/yyyy) is used for displaying and printing reports. Select **Sort dates ascending by default** if you want all documents to be sorted in ascending date order and set as the default.
Document Security	Select the security permissions for viewing documents in a case: • **If a document is in a non-accessible folder, it *is still accessible* in other folders a user can access**—(Default) Least restrictive: Allows users to view a document if the document is in a folder that they have permission to view (regardless of whether the same document exists in another folder that users do not have permission to view). • **If a document is in a non-accessible folder, it *is not accessible* in other folders a user can access**—Most restrictive: Prevents users from viewing a document if the document is in a folder that users do not have permission to view (regardless of whether the same document exists in another folder that users do have permission to view).

New Case: Processing Settings (Continued)

Field	Description
Tagging and Other Administrative Dates and Times	Specify whether dates and times are the same for case administration functions as for document display. Choose one of the following options: • **Use document dates and times**—Ensures that *all* date and time settings for the case (for administration and document display) are in the document format and time zone, as specified in the previous entry in this table. • **Use system dates and times**—Uses the system date and time settings for case administration tasks (such as user login tracking and export). Refer to *"Defining System Settings" in the -System Administration Guide* for information on the system level date and time settings. In the New York/London example, the administrator would choose **Use system dates and times** to keep administrative operations in the New York time zone (the system level time zone). However, if the all of the case administration and document handling were performed in London, the administrator would choose **Use document dates and times**.
Information Classification	
Enable automatic classification of incoming data. Note: Only policies enabled in the Information Classification portal will be utilized for classification.	Check to enable Information Classification in the platform. By default, Information Classification is disabled. **Note:** You must have enabled policies on the Information Classification portal side before enabling Information Classification in the eDiscovery platform. For more information, see *"Information Classification" on page 81*.
Modify search parameters	
Minimum size of document to return...	Enter the minimum size of documents to return when searching for documents with no indexed text: (default is 10 KB). **Note:** Changing this setting requires you to rerun post-processing.
Maximum result size (documents)	Enter the maximum number of documents (100 to 10,000,000) that can be retrieved by a search (default is 1,000,000).
By default, search filters show:	Specify the default view mode for filters. • Documents (family tagging) • Items
By default, search results show... attachments per document	Specify the number of attachments to show per message. The default is 10 and the choices are: 5, 10, 25, 50, and 100. **Note:** Only a limited number of attachments are shown in the search results until "show all" is clicked. Providing a boundary to the number of attachments per page helps overall performance.

New Case: Processing Settings (Continued)

Field	Description
Find Similar Settings	Select **Disable find similar in Review mode** if you want to disable default loading for Find Similar in Review Mode. If disabled, the "Find Similar" link will no longer display for reviewers when in review mode. Set the default document similarity threshold. This is the setting used in the similarity histogram as the default "Minimum Rating" value. A lower value indicates items which are less similar (versus a higher value indicating closer similarity, nearly duplicate) to the original item. **Note:** During review, users can adjust this similarity threshold for any original item to find similar items for analysis. For more information, refer to *"Viewing Search Results" in the Veritas eDiscovery Platform User's Guide*. You can also set where similar items are found: *across the entire case* or *within search results*.
Persistent Hit Highlighting	
Enable Persistent Hit highlighting for text	Under **Highlight Text**, enter a keyword or phrase. You can enter wildcards (* or ?) with single-word keywords only. Wildcards are not supported for phrases. You can also copy a list of keywords and phrases from a text editor and paste them into the **Highlight Text** field. The first 10 rows will be added as a separate **Highlight Text** field. Select the color to highlight the search term using **Highlight Color**. By default, the text is highlighted in Orange. When phrases are provided, only exact occurrences of that phrase will be highlighted, and any other individual occurrences of words part of that phrase will not be highlighted.
Enable Persistent Hit highlighting for Privacy Information patterns	Under **Hit Highlight Privacy Info**, select the required Privacy Info pattern from the **Privacy Info** list and select the highlight color. By default, the selected privacy info text will be highlighted in Pink. Note that you cannot repeat the selection of a privacy info pattern for persistent hit highlighting.
Note: You can select a maximum of 10 Highlight Text and/or Privacy Info fields for persistent hit highlighting. The system administrator can configure the maximum number of terms that can be entered for persistent hit highlighting by configuring the ***esa.queryengine.persistent.hit.highlight.terms.threshold*** property using **System** > **Support Features** > **Property Browser**. **Note:** Native viewer performance might be degraded if this property value is set to more than 20. The specified Highlight Text and Privacy Info patterns will be displayed and highlighted with the specified color in the documents for the selected case in the Document Review screen under Analysis & Review. You can also provide the Persistent Hit Highlighting settings for a case from the **All Cases** > **All Processing** > **Processing** > **Templates** tab while creating or modifying a case template.	

New Case: Processing Settings (Continued)

Field	Description
Define Active Directory parameters and specify internal domains **Note:** You cannot modify these settings after the case is created.	
Use Global Participants and Domains	If you use an Active Directory server to discover your Exchange servers and organizational data, you can modify this setting after the case has been created. However, once documents have been processed for the case, the setting is locked and cannot be changed. IMPORTANT: There may be distinct differences as to how participants and domains are resolved depending on whether this setting is checked or not. This setting may also affect participants, filter counts and search criteria. For more information, see *"AD Synchronization and "Use Global Participant and Domain" Case Parameter" on page 45*.
Internal Domains	To add a domain specific to this case, enter the domain name and click **Add**. To change a domain name, select the domain, enter the correct name, and click **Replace**. To delete a domain, click the trash icon for the name.

New Case: Processing Settings (Continued)

Field	Description
Specify text blocks to exclude from indexing	
Indexing exclusions	To exclude commonly found blocks of text from the index, enter the text on one or more lines, and click **Add**. To change a text block, select the text block, enter the correct text, and click **Replace**. To delete a text block, click the trash icon for the block. The specified text is excluded from documents processed in the future, but is not removed from the current index. **Note:** Spaces are ignored for disclaimer text identification. **Guidelines for specifying text block exclusions** • Disclaimer texts should be provided during case creation. • Up to 5 disclaimers each with a max of 8000 characters can be added. • During indexing disclaimer texts are identified and removed from indexing. • Disclaimer texts can be added/edited after case creation and indexing. However, changes will take effect only for new content indexed, so de-duplication can be impacted. • Disclaimer texts can be anywhere in the body of the email. • Disclaimer texts handling takes into account quoted texts, indentation and other special characters. • Disclaimer texts should match one or more lines in the body of the email. • Disclaimer text should end with the line. This means that no words should be found after the end of disclaimer text in the same line. For example: "I am a disclaimer. Please find me." – I am a disclaimer. Please find me. (will be matched) – I am a disclaimer. Please find me. If you can (will *not* be matched) • Disclaimer text does not need to be at the bottom of the email. • Disclaimer text block itself is filtered out of the email body and the rest of the data is indexed. • Disclaimer text provided should match a whole line in the body of the email or contiguous lines in the body of the email. Disclaimer texts cannot be just part of a line.

New Case: Processing Settings (Continued)

Field	Description
Configure processing parameters and features	
Estimated number of documents in index	Enter the estimated number of documents to be indexed (100,000 to 10,000,000). Used only to optimize performance (not a hard limit).
Messages with no Sender email address	Select one of the following: • Process and set sender to "none." Process the message and assign the value "none" to the **Sender** field. • Process and set sender to last modifier. Process the message and assign the email address of the last person who modified the email in the **Sender** field. • Do not process. Do not include the email in processing.
Enable Predictive Coding	Select the check box to enable predictive coding, (the ability to apply machine learning technology to learn the review criteria of your case and assess the corpus for relevant documents). For more information, see the Transparent Predictive Coding User Guide. **Note:** To enable predictive coding, you must also select the **Enable review, redaction, and production features** option under the **Enable/disable additional case features** section.
Process journal envelope information (Default is enabled)	By default, the system processes full Journal envelope messages. You can change the change the case settings prior to processing, to only process the original basic Journal messages (without envelope information). **Note:** Once processing has started, whatever setting you have chosen (on/off) cannot be changed. When enabled, the product will process, display in review, and provide searching capabilities for recipients who appear in the journal envelope. When disabled, the product will ignore the journal envelope, and will process the original message. **Considerations:** • For details on the supported versions of Enterprise Vault, see the *Identification and Collection Guide*. See *"Guidelines on Basic and Extended Journal Messages" on page 43*.

New Case: Processing Settings (Continued)

Field	Description
Enable discussion threads (Default is enabled)	By default, discussion threads are created during processing. When enabled, reviewers can see Discussions in search results after a case has been processed. Tag and bulk actions can be performed on a discussion thread, and discussion thread documents can be batched together. Reports will reflect discussion counts and discussions will be shown in participant, topic, attachment, and similar item analyses. This setting operates at the case level, so an administrator can decide during case creation whether faster processing is preferable to threading for a particular case. If you would rather process case data more quickly and users do not need discussion threads, uncheck this box during case creation. Once the first batch of data has been processed with discussion threads, the setting cannot be changed. If the setting is off, it can be turned on at any point. **Considerations**: Upon upgrade to version 8.2, existing cases will (usually) have **Enable discussion threads** on, whether or not any data has been processed. (The exception is if the last processing batch of a backed-up case did not run threader; the upgraded case will have **Enable discussion threads** off. If you re-enable discussion threads for a case, you will be prompted to re-run post processing.
Extract email files to (Default directory is given)	Specify the parent directory to which you want to extract contained PST and NSF files when found inside container files (such as ZIP files). This parent directory will contain a case specific folder (named for the case ID) when the case is created; this folder will ultimately contain the extracted files.
Convert supported mailbox files to PST	Enable, then specify which directory to place converted files. Setting this property overrides the system-level setting found at **System > Settings > Locations**. **Note:** The converted files directory is not included in the product's automated case backup. The default location places the directory in *<appliance_installation_drive>:CW\CaseData\<case ID>*. Ensure this location is a valid network share pathname (UNC) for any kind of environment that uses distributed processing (extracted email, distributed review or processing, etc.)
Crawler Properties for Non-Email Items	Setting the crawler properties for processing at a case level allows more granular control than at a system or global level. Select the non-email items (contacts, calendar items, tasks, journal entries and posts (files) from Exchange/PST, Notes/NSF and Archives to include in processing. **Note:** Email messages are always indexed for all document sources. These properties will be locked once processing begins.
Process loose files that are 0 bytes long	Select the check box to process files that are specified as 0 bytes in size.

New Case: Processing Settings (Continued)

Field	Description
Process truncated Notes documents	Select the check box to process NSF files that have been truncated by Lotus Notes and flagged by the system during the discovery process. The system processes the truncated Lotus Notes files and issues a warning. If you do not select the check box, the truncated Lotus Notes files are dropped from the source. You can exit the system and resolve the underlying issue for the truncation and then resubmit the NSF files for discovery and processing.
Document duplication in milliseconds	Selected by default, this option allows the product to de-duplicate documents based on the sent date of the document in milliseconds (rounded up to the nearest second). Clearing this check box means that documents will be de-duplicated, but only the seconds value will be used. **Note:** This applies to both loose files and e-mail, and can only be configured or modified prior to processing.
Interpret ambiguous "##/##/##"-formatted dates for derived emails as if formatted as	Select the date format for ambiguous dates (mm/dd/yyyy versus dd/mm/yyyy). A derived email is the text content of an email that is enclosed within another email. The product uses these emails to construct more complete and accurate discussion threads. However, because derived emails are text only, there can be ambiguities in how to interpret the sent date of the email.
Process a ".TIF" file's matching ".txt" file:	A TIF/TXT pairing is produced when documents are in imaged form (for example, scanned from paper documents). If optical character recognition (OCR) is applied to extract the text, the result is a pair of files that represents the content: an image (TIF format) and its extracted text (TXT format). The following options are supported. • Never. Process all ".TIF" files as regular image files, independent of matching ".txt" files. Do not perform any special actions when processing the file. • When the ".TIF" file is found in the specified folder and the matching ".txt" file is found in the specified folder. The system searches for a .txt text file that has the same name as the TIF file (such as "memo.tif" and "memo.txt") and is in the same folder. If the text file is found, it is processed instead of the TIF file. • When a pair is found within the same folder. The system searches for a .txt text file in the specified folder that has the same name as the TIF file in the other specified folder. If the text file is found, it is processed instead of the TIF file. • As described by a mapping file at the root of the source. The system searches for a text file that is mapped to a TIF file with the name that is found in the root folder of the source. If this mapping file is found and the corresponding text file is found, the text file is processed instead of the TIF file.

New Case: Processing Settings (Continued)

Field	Description
Container Extraction	Specify containers by type or extension to be extracted. Container extraction can be useful if you have a large dataset with a lot of containers and want to minimize any performance impact due to container handling at Discovery. Consult the "Not Processed" report to determine which set of containers to include/exclude during Discovery. See *"Generating Processing Reports" on page 99*. • Loose file containers are extracted during Discovery. • Attachment containers are extracted during Processing. These settings can be changed for each individual source. The various options are: • Maximum files in a loose file container: The maximum number of files within a loose file container that can be discovered. If the files within a container exceeds this number, the container is not extracted during Discovery. • Maximum files in an attachment container: The maximum number of files within a container attached to an email that can be processed. If the files within an attachment exceeds this number, the attachment is not indexed during Processing. • Limit Container Formats and Extensions for Loose files only. Default. Select this option to apply container extraction rules to loose files only. • Limit Container Formats and Extensions for Loose Files and Attachments. Select this option to apply container extraction rules to loose files as well as attachments. **Note:** Attachment container exclusions are permanent. Once excluded, these cannot be added back in future processing of the same source. • Container Formats: Select any or all of the format types from the list. • Container Extensions Exclude/Include Only <container extension>. This option excludes the extension provided if the format type is included in the Format Type options. Enabling exclusions based on format types can speed up the performance of both Discovery and Processing. – For example, to exclude a .JAR extension (which is a type of ZIP), include ZIP in the Container Format options. For strong file type and container exclusion workflow, see *"Guidelines on Container Extraction" on page 40*. **Note:** When attachment containers are excluded, metadata will still be created but content will not be indexed. Excluded attachments can also be retrieved for reviewing as well as for exporting. For Container File ID mapping information, see *"Supported Container Extraction File Types" on page 262*.

New Case: Processing Settings (Continued)

Field	Description
Processing Options	Specify criteria for processing: • All Dates/Dates On or After/Dates on or Before/Dates Between <Date> • All Sizes/Sizes Larger Than/Sizes Smaller Than/Sizes Between <size> • File Types: Select any or all of the file types from the list. • Exclude/Include File Extensions <list file extension>
Specify a filter to use when excluding known files	By default, the product uses the **NSRL Reference Data Set ("NIST" List)** to exclude known files during indexing. In addition to the default NIST list, custom lists can be defined in the System area. To add a filter to the menu, go to **System > Known Files**. **Note:** The selected list cannot be changed after indexing has begun.
Hidden, Inserted, and Embedded Content	By default, the product finds and indexes all text contained within a document. However, if the text is obscured or hidden, it can be difficult to find and view the indexed text. Identifying content enables you to search and filter for hidden and embedded content. Extracting embedded content enables you to view embedded documents as attachments or embedded content. • Selecting **Identify and Extract** option enables the following: – Identify all hidden content—(selected, but unavailable). – Extract all documents (for example, non-images)—(selected, but unavailable). **Note:** These options apply only to office and PDF loose files, and any attachments of these types in email messages from PST, MSG, EML/EMLX, NSF, and supported file sources. • Selecting one or more of the **Extract images from...** options to view embedded images in NSF documents, office and PDF files, and attachments. Optionally, select: • **Identify only**. Identifies whether there was embedded content, but does not extract the documents. • **Don't identify or extract**. Text is indexed even though the embedded content is not identified or extracted.
Audio Search	If licensed for audio search module, specify the **Default Language** to be associated with the case. There are 14 different languages to select from. • This setting can be overridden at the source level. • A source can be processed with only one language at a time. To correctly process audio files that contain multiple languages, make a copy for each language and process them separately. • If your system is not licensed for audio search or the audio search services are not running, the language selections will not display.

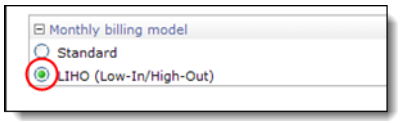
New Case: Processing Settings (Continued)

Field	Description
OCR Processing	
Use Optical Character Recognition (OCR) for documents where no text is found (image files, image-only PDFs)	By default, OCR is disabled. Selecting this check box enables OCR for documents where no text is found. Once a document has been processed with OCR, it cannot be re-done. For example, if you run OCR processing with one language dictionary, and later discover other languages in the batch, you will not be able to re-process those with the appropriate dictionary. **Note**: OCR processing will take much longer than normal processing. It is recommended that you not enable OCR initially. • Under **Apply OCR for:**, select all or specific file extensions that you want the product to product to process with OCR. If "PDF" is selected, PDF files will be recognized regardless of file extension. • Select minimum and maximum size **For files between:** to manage OCR processing. • **Language dictionaries to use.** Select among the various languages. English is the default but other supported languages are: Chinese (simplified), Chinese (traditional), Japanese, Korean, French, German, Icelandic, Italian, Portuguese, Russian, Spanish. The OCR engine will try to recognize characters from all selected dictionaries. **Notes:** • Processing case files requires more time when OCR is enabled. It is strongly recommended you leave this option disabled, with the exception of only very small cases. For normal size cases, leave this option off. Later, you can perform a search to select the documents you want to process with OCR. For more information, see *"Processing (or Resubmitting) Documents for OCR" on page 114*. • Version 7.0 and higher supports OCR processing of documents in Icelandic.
Languages **Note:** You can change all language settings after initial processing (except as indicated below in this table) and then rerun post-processing.	
Automatically identify the following languages within your case	Select check boxes to specify the languages that you want to include in document searches. Select only the languages that you believe may exist in your case. Languages that are not selected will not be automatically identified and will be classified based on the settings below. The most commonly-spoken languages are selected by default. **Note:** Version 7.0 and higher supports Icelandic.
When a portion of a document can be interpreted as more than one language	Sometimes the same words and characters are used in more than one language. This setting helps to accurately identify these shared words or characters. Specify the precedence order for determining the language (Chinese, Japanese, and Korean only). Click the **Move Up** or **Move Down** buttons to change the order.

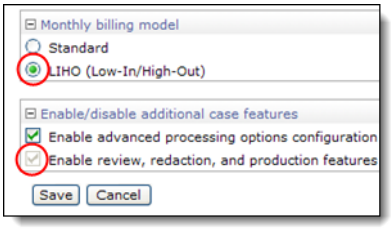
New Case: Processing Settings (Continued)

Field	Description
For documents that cannot be automatically identified	Select the single language to apply from the drop-down list if it is not possible to identify languages in a document automatically. For example, it is difficult to accurately identify documents with limited content, such as short emails and appointments. If the expectation is that your data set is mostly in one language, such as English, then configure this setting to that language to best classify these documents. Alternatively, you can classify these documents as "Other."
Advanced Options	For small amounts of document content, it is not possible or desirable to automatically identify the language. You can configure the minimum number of characters and the percentage of a document's content that is required to automatically identify a language within the document. Exceeding either the character or percentage threshold will trigger automatic language identification. When you click the **Advanced Options** button, the Automatic Language Identification Advanced Options window opens. Configure the following settings: • Specify the minimum number of characters to automatically identify a language (default is 200). • Specify the minimum percentage of a document's content to automatically identify a language (default is 10%). • For content that does not meet the other thresholds or cannot be automatically identified for any other reason, choose a language for manual identification.
Enable stemmed search for the following languages	Select check boxes to enable stemmed searches for specific languages. A stemmed search automatically finds documents that contain common variations of a word that is entered as part of a query. For example, if you search for the word "test," a stemmed search also finds variations such as "testing,""tests," and "tested." Two English options are available to support stemmed searches. Both are selected by default: • English—Uses a sophisticated linguistic stemming algorithm to determine stemming rules. For example, this option considers "went" as a variant of "go." • English (suffix-based stemming)—Uses the Porter algorithm to strip out common word suffixes (such as "s" or "ing") for stemming. This algorithm is useful for finding nouns in their plural and singular forms. **Note:** Each additional language increases processing time within your case.
Monthly Billing Model	

New Case: Processing Settings (Continued)

Field	Description
Standard or LIHO (Low-In/High-Out)	By default, the product is enabled for Standard billing. However, if you prefer to be billed less for processing documents in a case, and be charged only for the documents selected for review and/or export, select LIHO as your monthly billing model. **Note:** You must first have a consumption based license in order to view and select the LIHO billing option. 

New Case: Processing Settings (Continued)

Field	Description
Enable/Disable Licensed Features	
Enable advanced processing options configuration (also known as pre-processing)	Enable or disable the options for document pre-processing. (This option is available only if the appliance is licensed for processing options.)  **Note:** Selecting LIHO as your billing model automatically enables advanced processing options and review for your case and cannot be changed. (Documents must be in a designated review folder to be flagged for billing purposes.) **Note:** If you do not have a license for the Pre-Processing module, or if the module is disabled at case setup, you will not be able to process LEF files, de-NIST loose files, or get Sent dates in email files (PST, MSG, EML, NSF).
Enable review, redaction, and production features	Enable or disable options for document review, redaction, and production. (Available only if the appliance is licensed for these features.) **Note:** This feature must be enabled for Predictive Coding to function.
Enable email header viewer	Enables the viewing of email headers in email messages. By default, this option is disabled.

New Case: Processing Settings (Continued)

Field	Description
Enable email header search	Enables the searching of email headers in email messages. By default, this option is disabled. **Note:** This option is not supported for pre-7.1.3 upgraded cases. **Caution:** Indexing email header search content can be a resource-intensive task. Be aware that enabling email header search may degrade system performance. The extent to which enabling the email header search option affects system performance depends on the size and composition of your case data and the email header fields selected for indexing. Unless you have identified email header fields (for example, Standard or Custom email header fields) as useful search criteria, leave this option unchecked. • **Enter Email Headers Field for Indexing**. You have the choice of adding, deleting, or not editing the list of predefined **Email Headers.** If you need to analyze and examine email header fields that are not part of the pre-selected (default) choices, the product provides the ability to add custom metadata email header fields (provided the fields and values are RFC 822/2822 standard compliant) in the text box. Your selections display in the **Index the following email header fields** menu. See *"Date Formats for Email Header Fields: Received and X-Received" on page 38*. **Considerations:** • You can delete or change email header field entries in the case settings window but once processing has started, whatever has been entered cannot be changed. • Since all standard fields (To, Cc, Bcc, From, Sender, Sent, Subject, Date, Importance, Priority, Sensitivity, etc.) are already added to the system by default, the system prevents you from adding them again. An error message is displayed to inform you of this issue. • **Email Header Fields are case-insensitive**: The product supports case-insensitivity for email header fields. Upper- and lower-case differences in email header fields are ignored and all fields are converted to lowercase before indexing. This means that "APPROVED-BY", "approved-by", ApproVed-By" and "Approved-By" are all matched for indexing and for search results. **Note:** You may want to refer to the RFC standards document (http://tools.ietf.org/html/rfc2822) for information on the specific syntax of RFC 2822 messages and to view helpful examples.

Date Formats for Email Header Fields: Received and X-Received

The following table lists date formats that you can use with the Email Header fields: `Received` and `X-Received`.

.

Received and X-Received Date Formats

	Date Pattern	Example
1	EEEEEEEE, MMMMMMM dd, yyyy	Tuesday, March 10, 2013
2	yyyy MMMMMMM dd	2013 March 10
3	dd MMMMMMM yyyy	10 March 2013
4	MMMMMMM dd, yyyy	March 10, 2013
5	EEE MM/dd/yy	Tuesday 04/10/13
6	EEE dd-MMM-yy	Tuesday 10-Mar-2013
7	EEE dd/MM/yy	Tuesday 10/04/13
8	MM/dd/yy	04/10/13
9	dd/MM/yy	10/04/13
10	EEEEEEEE, dd. MMMMMMM yyyy	Tuesday, 10.March 2013
11	MM/dd/yy	04/10/13
12	dd/MM/yyyy	10/04/2013

Date Formatting Notes

You can add additional date formats that are not included in the Date Format table above with the Property Browser.

Note: If a date in a header field is not in the list of date formats listed below, it *is* indexed, but it is not converted and stored as a date, so only exact matches will find it.

To add date/time formats for Email Header Field: Receive

The example below adds two date formats with hyphen separators.

1. Go to **System > Support Features**, and select **Property Browser**.
2. Enter the property: `esa.indexer.emailheader.dateformat.13`
3. Set the value to the new date format: `dd-MM-yyyy`
4. Click **Submit** to save your date/time setting.
5. Repeat Step 2 and increment the count (for example, ".14").
6. Enter the property: `esa.indexer.emailheader.dateformat.14`
7. Set the value to the new date format: `MM-dd-yyyy`
8. Click **Submit** to save your date setting.

Guidelines on Container Extraction

In container extraction, container files (such as ZIP and RAR files) are examined and have their contained files extracted and processed as individual files for analysis, review, and production.

Contained files must be separated from their containers because different files within the same container may have a different status (such as relevant or privileged) and must be handled separately from their companion files.

The container extraction option **Limit Format Types and Extensions for Loose files only** is enabled by default for all new cases.

If you choose not to perform container extraction, then container file text is still fully searchable, but the container will be processed as a single unit and will appear as a single container file for search, review, and export. Further, when container extraction is disabled, and the system finds a container in EML, MSG, PST or NSF documents, the files will NOT be expanded. However, note that for loose files, container extraction is always enabled.

The following rules apply to container extraction.

- Container extraction is supported only for Releases 4.0 and later.
- The default number of supported container files for loose files is 10,000.
- The default number of supported container files for attachments is 1,000.
- When the system encounters a container file in a loose file directory (not attached to an email), it extracts all the contained files from within the container and processes them as individual documents. It does NOT process the container file itself as a document. Instead, the container file is treated as an element of each contained file's path and is available for viewing search through the document locator.

  For example, if the file *mydocs.zip* contains the files *budget.xls* and *memo.doc*, these two documents are added to the index. There will be no document entry for *mydocs.zip*; however, when viewing the contained files an icon is displayed indicating that it was found in a container, and the container filename itself can be viewed and searched on through the document locator. The Case Status screen will show that one container file was encountered during processing, and that two files were extracted.
- When container files are email attachments, they are treated similarly to loose files. The contained files are extracted and shown as individual attachments on the email, but the container itself is not shown as an attachment. However, the product will show the name of the container file in a hierarchy in the attachment view (on both the message and in the related items area) to make it clear which attachments came from container files. On export, no specific document entry will be included for the container, but the name of the container file will appear in the locator path for all of its child documents.
- If you have a container attachment that has some responsive documents and some non-responsive documents, the whole email must be tagged consistently as if the email had regular, non-contained attachments that were responsive and non-responsive.
- ZIP, GZIP, RAR, TAR, LZH, LHA, Unix compressed, BZ2, and 7Zip container files are supported, as well as self-extracting (.exe) ZIP and RAR files. Other unsupported container files are passed unaltered to the indexer.

- If the system cannot open a container file due to password protection, encryption, or other reason, the container file is dropped and logged as an error. If there is a single encrypted file within a container, the full container is not processed.
- If the system encounters a problem extracting a specific file from a container (other than encryption), its content is not indexed, and a file warning will be logged. The path of the dropped file will include the name of the container file and any relative path information. While the content is not indexed, a document will be created and the generic file information will be indexed. An error may occur when exporting the file similar to the error encountered during processing.
- Containers with up to 10,000 files are supported by default. You can increase this limit by modifying the number on the case setting page under **Processing > Settings > Configure processing parameters and features > Maximum files in a loose file container:**.
- You can increase this limit by modifying the property *esa.indexer.max.docs.percontainer*. If a container has more than the allowed maximum number of files, then the entire container file is dropped. There is also a timeout threshold of a few minutes for extraction.
- During export, contained files are treated as individual files, but a reference to the container is maintained. For example, if the file *memo.doc* was exported from container file *mydocs.zip*, the export directory structure includes the file: `..\mydocs.zip\memo.doc`. From email attachments, the email is exported as a single original unit with the original container files.
- In XML exports, contained files are exported as individual files with the reference to the container file in the document location record of the XML metadata. For email attachments, if the original native email is exported then the original container file is preserved in its original format. However, if the **Separate attachments from emails** option is chosen, the contained files are exported as individual attachments with the reference to the container file in the document location record of the XML metadata.
- If a PST or NSF file is encountered in a loose file container, it will be extracted to the case's PST/NSF extraction area and processed like any other PST or NSF file. However, PST/NSF files found within attachment containers will not be extracted and will instead be flagged with a warning and logged in the case's **Exceptions** screen **> File Notices** tab.
- MSG and EML files found within loose file containers are processed as emails. MSG and EML files that are attached to other emails are processed as loose files. However, all of their child documents (attachments/additional embedded messages) will be broken out and indexed as separate attachments during indexing, and the full attachment hierarchy will be displayed and preserved on export.
- The loose files containers that were excluded based on the options selected can be selected to be extracted in the next batch by changing the container-extraction options and re-running discovery on the same source. For attachment containers, once they are excluded, they cannot be extracted in the same source by re-running processing.

- Strong File Type and Container Exclusion Workflow

  There are 3 ways to exclude container files during discovery. You can exclude containers based on:

  – Size (number of files)

  – File extension

  – Strong file type

  Excluding container files uses strong file types and follows a different workflow than container size or extension. Follow these steps:

  – Run Discovery excluding all container formats (no **Container Formats** selected).

  – Run the "Not Processed Documents" report. See *"Generating Processing Reports" on page 99*.

  – Review the report to cross-check and compare container extensions with their strong file type counterpart. From the report output, determine which container formats you want to process (include) or exclude.

  – From the **Processing > Sources and Pre-Processing > Pre-Processing Options > Case Folders** screen, select the container formats for processing.

  – Navigate back to **Sources and Pre-Processing > Pre-Processing Options > Manage Sources** screen and check the appropriate source.

  – From the **For selected items** drop-down menu, select either **Discover new files for source** or **Start processing source with discovery** and click the **Go** button.

  Note: Because container exclusion utilizes strong file typing, any container file extensions that do not match their underlying correct (strong) file type, are processed according to their strong file type. For example, a .ZIP that, for whatever reason, is actually a .RAR container file is treated as a .RAR by the system.

  For Container File ID mapping information, see *"Supported Container Extraction File Types" on page 262*.

Guidelines on Basic and Extended Journal Messages

When configured in case settings, the product adds envelope information to the original Journal message. It may be helpful to think of the Journal message as a child message containing the regular "To", "Cc" and "Bcc" fields and Journal Envelope as a parent message which has these fields merged with its own equivalent fields. This additional information allows users to view, search, and export expanded distribution lists for greater accuracy in filtering on, and producing what was actually indexed in the system.

MSG and EML Formats

The platform supports both MSG and EML formats. The difference in the formats is related to the two ways in which the Exchange server is set up for journaling. The first way is journaling to an Exchange mailbox, and the second way involves journaling to an SMTP server. Exchange mailbox means the format is MSG and SMTP means the format is EML.

Basic and Extended Journal Messages

There are two formats in which the product treats archived Journal messages:

- Extended Format (*Exchange 2007 and later*)—If the Journal message that was archived was in extended format, the product retrieves the following fields from EV: Journal "To", Journal "Cc", Journal "Bcc" and "Recipients". In this case, the regular fields of the child message are merged with the corresponding parent fields.
- Basic Format (*Exchange 2003*)—If the Journal that was archived was in basic format, the product retrieves only the "Recipients" list (without further classification of the other three fields). In this case, the recipients list is added to the regular "To" field.

Process Journal Envelope Information: Enabled versus Disabled

When enabled, the product will process, display in review, and provide searching capabilities for recipients who appear in the Journal envelope. When disabled, the system ignores the Journal envelope, and processes only the original message.

Deduplication

Deduplication is done at the parent level. For example, a child email message is not deduplicated against an email which has both parent and child information. All copies of a parent email are run through deduplication. For example, if several employees are being journaled to different Journal mailboxes, the system collects all the Journal mailboxes. If multiple employees were sent the same email, those email messages will be deduplicated.

To disable Journal envelope processing:

1. Either create a new case, or select an existing case (not yet processed) and from **Processing**, click **Settings**.
2. Click to expand the *Configure Processing Parameters and Features* section, then clear (deselect) the **Process Journal Envelope information** option.
3. Enter/change any other information, then click **Save** to save your settings.

Additional Notes

The following notes also apply to this Journal Envelope message feature:

- Available for new collections only.
- Existing collections cannot be converted (as Journal data in the PST files would not be available).
- MSG journal support is only available for cases created on v8.0 onward. EML journal support is only available for cases created in v8.1.1 onward. If your case was originally created on earlier versions of the product this functionality will not be available.

The product retrieves the expanded distribution list members from EV (from the Journal metadata), not from the Custodians list.

Discovering Archive Sources

Note: You must have the **System Manager** role to perform all system configuration and archive discovery tasks on the appliance.

For information about how to discover and manage archive document sources, refer to the following topics:

- *"About Active Directory Discovery" in the next section*
- *"About Discovering Veritas Enterprise Vault (EV) Sources" on page 48*
- *"About Discovering Lotus Domino Sources" on page 48*

About Active Directory Discovery

The Active Directory (AD) crawler discovers your Microsoft Exchange servers, the mailboxes on each server, and your organizational data, such as physical locations and departments (groups). The appliance must belong to a Windows domain for the AD crawler to run. To schedule the AD discovery to be run periodically, refer to *"Managing Schedules" in the System Administration Guide*.

To index the documents on a discovered Exchange server, the server must be added to a case (see *"Defining New Cases" on page 18*).

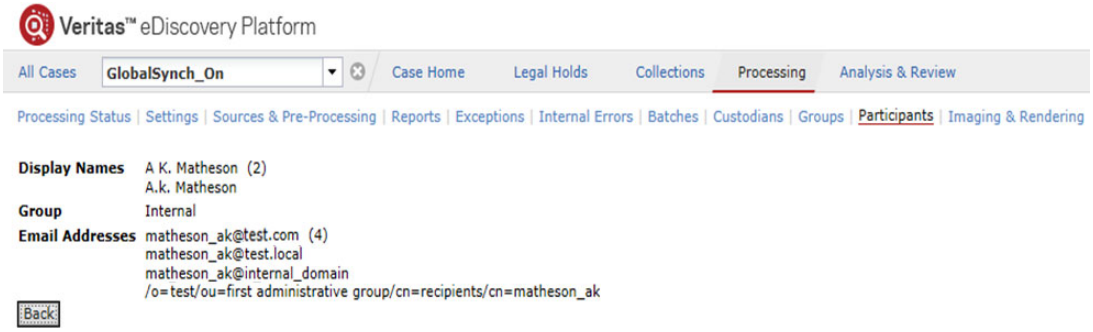
AD Synchronization and "Use Global Participant and Domain" Case Parameter

The "Use Global Participant and Domain" case setting affects the way participant emails and domains are resolved. This parameter can also influence filter and search results. It is a good idea to review the possible outcomes and results discussed in this section.

During initial setup, AD discovery is run to synchronize AD domains and users (participants). For subsequent batches of data that are ingested into a case where the case setting **"Use Global Participant and Domain"** was enabled, any additional or incremental participant information is automatically updated for the case. The platform uses this information to resolve email addresses (domains) that are present in the **FROM**, **TO**, **CC** and **BCC** email fields.

Note: The setting **"Use Global Participant and Domain"** is located in **Processing** > **Settings**, under section **Define Active Dictionary parameters and specify internal domains**. The default setting is disabled.

In the following case example, GlobalSynch_On, the case setting, **"Use Global Participant and Domain",** is enabled and the participant A.K. Matheson is part of the internal group domain and is associated with 4 email addresses.



Conversely, if AD sync was *not* enabled in the case setting, the platform's first attempt is to try and correlate multiple email addresses to the same person. However, there is a possibility that it will not be able to relate and resolve the email addresses. In this case, the 4 emails with the above email addresses are not mapped back to A.K. Matheson. Instead, they would be resolved and treated as four separate participants. In this latter scenario, participant searches as well as filter counts will reflect this behavior and would not correlate that the 4 mails addresses belong to 1 participant.

Case Differences Between AD Synchronization Enable/Disable Setting

Item	AD Sync Enabled	Ad Sync Disabled
Participant search will resolve all email addresses belonging to a participant. See example above.	Yes	No
Participant count and details in **Case > Processing > Participants**	Full listing of all associated participants and addresses listed in FROM, TO, CC and BCC fields.	Display of participant but not any associated participants.

Case Differences Between AD Synchronization Enable/Disable Setting (Continued)

Item	AD Sync Enabled	Ad Sync Disabled
Domain Default	Exchange-style addresses are all marked as **internal** domain.	If there is no active directory information when AD synchronization is not enabled, then participants are put into an **external** domain group. If this is not correct, you must explicitly designate the domain as internal during the domain list setup in order for participants to be included in the internal domain group.
Participant and filter counts in Analysis & Review	Counts include all other email addresses that the participant has and will indicate the **internal** domain.	Counts mapped to Participants and domains will differ for all non-exchange-style email addresses. The mapping and correlation of email relationships are not maintained.

For details on how to perform Active Directory discovery, refer to *"Setting up Data Sources" in the Identification and Collection Guide*.

About Discovering Veritas Enterprise Vault (EV) Sources

For detailed information on performing EV discovery, refer to *"About Enterprise Vault Discovery" in the Identification and Collection Guide*.

To schedule the Veritas EV discovery to be run periodically, refer to *"Managing Schedules" in the System Administration Guide*. To limit the appliances that can access a discovered archive, refer to *"Managing Schedules and Jobs" in the System Administration Guide*. To index the documents on a discovered vault, the vault must be added to a case (see *"Defining New Cases" on page 18*).

For details on the supported versions of Enterprise Vault, see the *Identification and Collection Guide*.

About Discovering Lotus Domino Sources

To schedule discovery on the Lotus Domino source to be run periodically, refer to *"Managing Schedules" in the System Administration Guide*. To limit the appliances that can access a discovered archive, refer to *"Managing Schedules and Jobs" in the System Administration Guide*. To index the documents on a discovered vault, the vault must be added to a case (see *"Defining New Cases" on page 18*).

For detailed information on performing discovery on Lotus Domino Sources, refer to *"Lotus Domino® Server Setup" in the Identification and Collection Guide*.

Managing Case Sources and Custodians

For information about how to manage case document sources and custodians, refer to the following topics:

- *"Selecting Document Sources and Pre-Processing" in the next section*
- *"Processing Physical Evidence Files (LEF and E01)" on page 59*
- *"Defining Case Custodians" on page 62*
- *"Merging Custodians" on page 63*
- *"Unmerging Custodians" on page 65*
- *"Assigning Custodians" on page 66*

Note: Refer to the *Load File Import Guide* for adding load file sources.

Selecting Document Sources and Pre-Processing

For each case, you can index documents from any combination of the following sources:

- Selected loose files, email container files (Microsoft Exchange PST or Lotus NSF), or individual email files (.msg or .eml).
- Selected Exchange mailboxes, archives, and/or repositories on the discovered email server/ archive sources (Microsoft Exchange, Veritas Enterprise Vault)
- Selected collection sets and load files. Refer to the *Load File Import Guide* and *Identification and Collection Guide* for more information.

Indexing can be run manually for each source, or you can schedule a document crawler task to periodically update the case index with any new content found in one or more sources (see *"Managing Case Schedules" on page 204*).

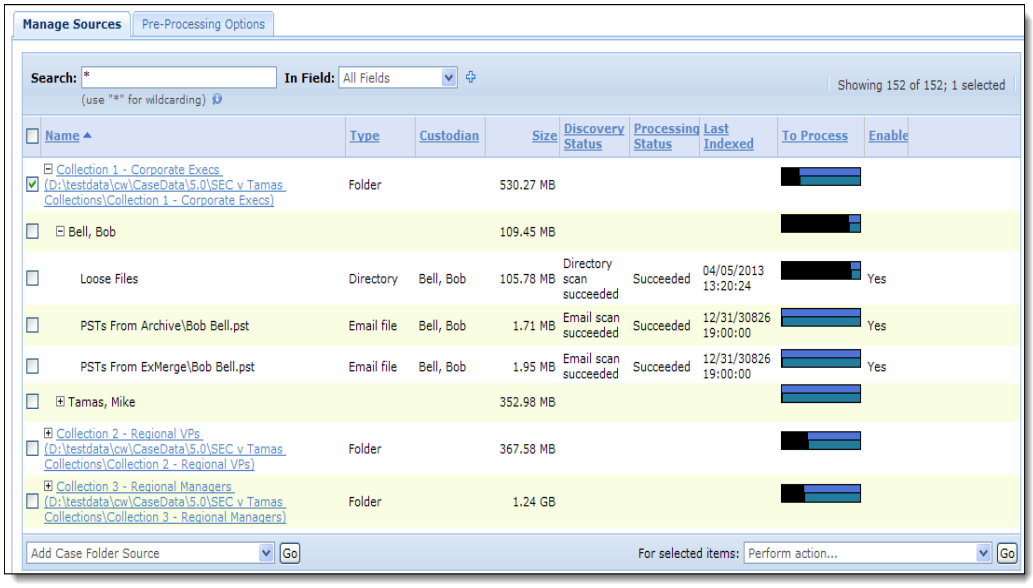
Note: The speed at which documents are processed into the product can vary dramatically based on the type of content and the configuration of your document sources and network. Contact Technical Support for information about optimizing your setup to maximize document processing throughput.

Defining Document Sources

To define the document sources for a case

1. On the top navigation bar, select a case, then click **Processing > Sources and Pre-Processing**.

   The screen opens to the **Manage Sources** tab.



2. To search the list of mailboxes, files, and directories:
   - In the **Search** field, enter the starting characters of the text to be found (use a "*" to indicate any text). For example, to find all names that start with "Robert" enter "rob" or "*ob".
   - From the **In Field** menu, select the column to be searched.
   - To add additional values for searching, click the plus icon. The search finds any of the entered values (OR search). Click the minus icon to remove the additional values.

For each source, the screen includes the following information and controls.

Manage Sources screen Columns

Column	Description
Selection check box	Check box to select the entry.
Name	Name of the source. To expand an entry and show the associated directories, click the + sign to the left of the entry. Click the - sign to collapse the entry.
Type	Folder, directory, or email file.
Custodian	Custodian name, if a custodian is assigned.
Size	Size of the email file (PST or NSF).
Discovery Status	Status of the last discovery job.
Processing Status	Status of the last processing job.
Last Indexed	For a folder, directory, or file the column lists the time it was last indexed. For PST files, this is the last crawl date, as specified in processing options. If no dates were specified for PST, the product crawls to the date 30826 and processes. For NSF files, the Last Indexed time (if no dates were specified in the processing options) is the machine time when the last indexing was run on this source.
To Process	Status of processing options as defined on the **Processing Options** tag. See *"Pre-Process Your Source Data" on page 67*.
Enabled	Indication of whether the entry is enabled for indexing.

3. To perform actions on the selected sources, select the check boxes for the sources. Choose one of the following source options from the menu in the lower-left, or choose an action from the menu in the lower-right corner of the screen and click **Go**.

   Note: You can choose to apply an action only to specific rows (such as an entire source, or one or more email files). If you attempt to perform an action on a row that is not permitted, a message indicates how the action should be performed.

Source Options on Manage Sources screen

Action	Description
Add Case Folder Source	Add a new document source to the case. See *"Adding Case Folder Sources" on page 53*.
Add Load File Source	Add a new third party load file source to the case. Refer to the Load File Import Guide.
Add Collection Set	For information on how to add a collection set, refer to the section *"Processing Collection Sets" in the Identification and Collection Guide*.

Source Options on Manage Sources screen (Continued)

Action	Description
Rerun Post-Processing	Apply changes that you have made to this screen. **Note:** This applies when merging custodians, assigning processed data to new custodians, or changing language options for the case. This also applies if cases have just been upgraded, or have stopped processing jobs, in which some data has been processed. A warning displays when post-processing should be re-run for the specified source (not for all source changes made to this screen).
Export Table	Export the source list in CSV format.
View Exceptions	Open the Exceptions screen for the selected case. See *"Monitoring Source Processing Status" on page 107*.
Show All Sizes in GB/Show Sizes in KB/MB/GB	Change the document units shown on the screen to be all in GB or in KB, MB, or GB, as appropriate according to the file size.

Actions for Selected Items on Manage Sources screen

Action	Description
Discover new files for a source	Search the specified sources for new email files to index. **Note:** Search applies only to selected item(s).
Check email file integrity	Scan email files to verify integrity prior to processing. This allows you to ensure that email files are free of corruption and can be properly processed. For example, if a scanned email file is found to have issues, the system automatically disables the email file so you can repair it. After the file is repaired, you can rescan it. If the rescan is successful, the file is re-enabled for processing.
Start processing source with discovery	Discover any newly added data to the case folder source, and start the indexing process. The system prompts you to specify an optional batch processing label. The label is used in the Manage Batches area and on each indexed document. See *"Managing Batches" on page 144*. You can monitor the status of ongoing processing jobs through the Jobs window. **Note:** If you have already performed discovery and/or pre-processing analysis on your sources, and no new files have been added, then do not use this option. Save time by choosing **Start processing source without discovery** instead.
Start processing source without discovery	Start the indexing process and do not search for new files to index.
Stop processing source	Stop the indexing process.

Actions for Selected Items on Manage Sources screen (Continued)

Action	Description
Set Processing Options	Set processing options to apply to only this source. When you select this option and click **Go**, a pop-up window opens. Configure the following settings and click **Go**: • **Date**—Select a date option and use the calendar icon to specify the dates. • **Size**—Select a document size option and specify the size range. • **Document Types**—Select check boxes for the document types that you want to include in processing. To select or deselect all of the document types, check or clear the check box at the top of the list. • **File Extensions**—Enter the file extensions of files to exclude from indexing, such as EXE and DLL files. Use a space or comma to separate multiple entries. These values apply to loose files only, not to email attachments. All email attachments are processed regardless of the file exclusion list. **Note:** See *"Pre-Process Your Source Data" on page 67* for more information.
Enable processing	Activate processing.
Disable processing	Deactivate processing.
None	Do not assign a custodian.
New custodian	Assign a new custodian to a source or sources. When you click **Go**, the system prompts you for the name of the new custodian. When you click **OK**, the new custodian is created and assigned to the source or sources.
Custodians	Select the custodian name to assign the custodian to that subsource. Custodian assignments take effect for the next processing or post-processing run. For more information, see *"Defining Case Custodians" on page 62*.

Adding Case Folder Sources

You can point to a single directory to automatically process all loose files and emails within the directory. The following rules apply:

- You can have up to three discovery jobs, and one processing job to be active on an appliance at one time (with cases starting in 6.0 or later). Across cases, the product can have as many other discovery and processing jobs running simultaneously (as determined by memory availability and CPU, starting with version 5.5).
- In order to process documents for multiple cases simultaneously, it is necessary to create those cases on different nodes of a cluster. The node that a case is created on can be specified on the Configure Case screen. This setting cannot be changed after the case is created; however, it is possible to move a case from one node to another through the backup/restore process.
- The file discovery scanning rate can vary depending upon data type.

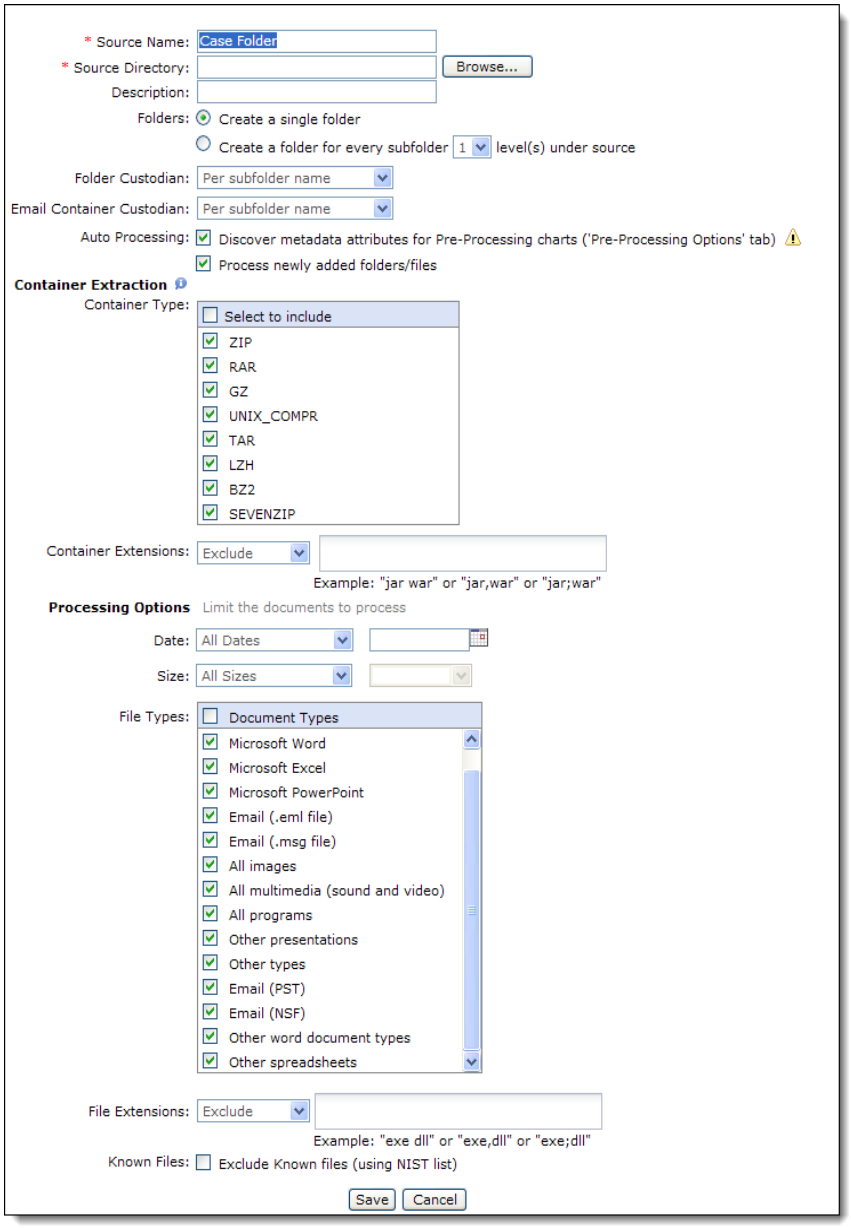
You can use the Add Case Folder Source screen to add the documents (email files and loose files) for a case.

Note: If you will be processing the same documents into multiple cases, you must create a separate physical copy of the files for each case and create a case folder.

To add sources to a case

1. On the top navigation bar, select a case, then click **Processing > Sources and Pre-Processing**.
2. On the Manage Sources screen, select **Add Case Folder Source** from the menu in the lower-left corner of the screen, and click **Go**.

   Note: Some of the settings that are available on the **Add Case Folder Source** menu depend on what you select as case settings. For example, if you select the case setting "Enable advanced processing options configuration (also known as pre-processing)", then settings on the **Add Case Folder Source** page under the Processing Options section include File Types.


* Source Name:
Case Folder
* Source Directory:
Browse...
Description:
Folders:
Create a single folder
Create a folder for every subfolder 1 level(s) under source
Folder Custodian:
Per subfolder name
Email Container Custodian:
Per subfolder name
Auto Processing:
Discover metadata attributes for Pre-Processing charts ('Pre-Processing Options' tab)
Process newly added folders/files
Container Extraction
Container Type:
Select to include
ZIP
RAR
GZ
UNIX_COMPR
TAR
LZH
BZ2
SEVENZIP
Container Extensions:
Exclude
Example: "jar war" or "jar,war" or "jar;war"
Processing Options
Limit the documents to process
Date:
All Dates
Size:
All Sizes
File Types:
Document Types
Microsoft Word
Microsoft Excel
Microsoft PowerPoint
Email (.eml file)
Email (.msg file)
All images
All multimedia (sound and video)
All programs
Other presentations
Other types
Email (PST)
Email (NSF)
Other word document types
Other spreadsheets
File Extensions:
Exclude
Example: "exe dll" or "exe,dll" or "exe;dll"
Known Files:
Exclude Known files (using NIST list)
Save
Cancel

3. Enter the following information. An asterisk (*) indicates a required field.

Case Folder Information

Field	Description
Source Name*	Enter a name for this source (up to 255 characters). Use only letters, numbers, and underscores. The name should help identify the type of source, such as "Atlanta Collection."
Source Directory*	Click **Browse** and select the top level folder for the case on the appliance or enter a remote directory name, click **Go**, and select the appropriate folder. Click **OK**. Your network access depends on the Windows name and password specified in the system settings under Indexing (refer to *"Defining System Settings" in the -System Administration Guide*). Alternatively, enter the full path of the source directory in Uniform Naming Convention (UNC) format (up to 256 characters). For example, if a PST folder is on a remote device: *\\pine\pstfolder* If the folder resides on the appliance: *C:\PSTFiles*
Description	Enter a description of the source.
Folders	Select the folder level appropriate for this source: • Create a single folder. Add all documents to a single folder. • Create a folder for every subfolder. Create a new folder for each subfolder in the original source tree. Include only the levels of interest. **Note:** When you point to subfolders within a case folder directory, the system does not process any files that are found at higher levels. To check that your case folder setup is accurate, you can obtain the document count in Windows Explorer at the case folder level and make sure that the count matches the file count on the Case Status screen.
Folder Custodian	Custodians allow users to search for case documents according to the individual identified as responsible for the documents. Select a default custodian associated with all files discovered in the source directory in one of the following ways: • To use no custodian, select **None**. • To define a new custodian, select **New custodian**, enter a custodian name, and click **OK**. • To assign the custodian with the same name as a subfolder name, select **Per subfolder name**. This is a convenient way to assign custodians to folders. Use the custodian name as the folder name, and then select this option. • To select a specific custodian, choose the custodian from the menu. Example: The directory structure is *C:\my case documents*, with the files ..*Custodian 1* and ..*Custodian 2*. If you select a level of "1" and set the folder/email custodians to the folder name, all emails/files under "Custodian 1" will be assigned the custodian "Custodian 1." To override the default custodian for specific files, see *"Defining Case Custodians" on page 62*.

Case Folder Information (Continued)

Field	Description
Email Container Custodian	Select a default custodian associated with all emails containers discovered in the source directory in one of the following ways: • To use no custodian, select **None**. • To define a new custodian, select **<New custodian...>** enter a custodian name, and click **OK**. • To assign the custodian with the same name as a subfolder name, select **Per subfolder name**. This is a convenient way to assign custodians to folders. Use the custodian name as the folder name, and then select this option. • To select a specific custodian, choose the custodian from the menu. To override the default custodian for specific files, see *"Defining Case Custodians" on page 62*.
Auto Processing	Select one or both check boxes to specify whether the following will be discovered/processed automatically: • Discover metadata attributes for Preprocessing charts **Note:** This option applies to loose files, and MSG/EML, PST, and NSF files. • Process newly added folders/files
Container Extraction	Select check boxes for the container formats that you want to include in processing. To select or deselect all of the container formats, check or clear the check box at the top of the list.
Container Extensions to Exclude	Enter the container extensions of files to exclude from indexing, such as "JAR WAR". Use a space or comma to separate multiple entries.
Processing Options	Specify the date and time range for indexing the source files. For loose files, the range applies to the last modified date/time and for email files it applies to the sent date/time. • Click calendar  icon, enter the time in 24-hour format, and select a month and day. or • Enter the date and time directly as: MM/DD/YYYY HH:MM:SS. **Notes:** – The date/time restrictions do not apply to new files that are added to directories that have already been indexed. To use the date/time restrictions, place new files to be indexed into new directories.
Document Types	These settings are visible only if the pre-processing module is included. Select check boxes for the document types that you want to include in processing. To select or deselect all of the document types, check or clear the check box at the top of the list.
File Extensions to Exclude	Enter the file extensions of files to exclude from indexing, such as EXE and DLL files. Use a space or comma to separate multiple entries. These values apply to loose files only, not to email attachments. All email attachments are processed regardless of the file exclusion list.

Case Folder Information (Continued)

Field	Description
Check integrity of newly added email files	Select the check box to automatically verify the integrity of email files that prior to indexing.
Process newly added folders/files	Select the check box to automatically index all newly added folder and files.

4. Click **Save** to save the new source, or click **Cancel** to discard your changes.

Processing Physical Evidence Files (LEF and E01)

Note: To process any of Guidance's forensic imaging formats, load the file within Encase and convert it to a logical evidence file (LEF) or an E01 file. If you create an E01 file, ensure that you create an MDM file as well. For more information, see *"LEF" on page 237*.

LEF Files

LEF files are processed directly.

To add an LEF file

- Place the LEF in a folder and add the folder as a source.

  It will be processed like any PST, NSF, or loose file.

E01 Files

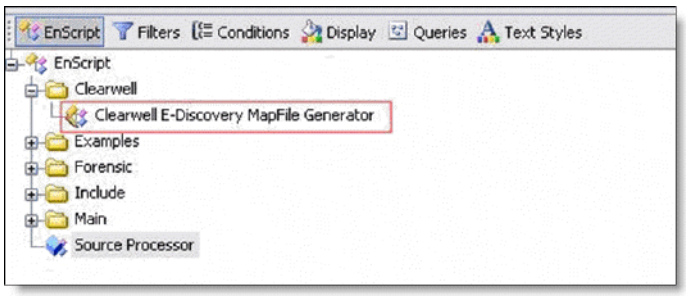
A special process is required to prepare physical evidence files (E01s) for processing as part of a case folder source, because E01 files do not include readily accessible metadata. To extract the metadata that is required for processing, you must first process the E01s using the eDiscovery Mapfile Generator. The setup file (setup.exe) for eDiscovery Mapfile Generator is available at ***\CW\<version>\utilities\EncaseEnscript***. For example, ***D:\CW\V95\utilities\EncaseEnscript***.

Note: The eDiscovery Mapfile Generator is only supported on the 32-bit version of Encase. To use the eDiscovery Mapfile Generator, you must first install the Encase software. It is important that the Encase software and the EncaseEnscript utility are installed on a different server than the appliance.

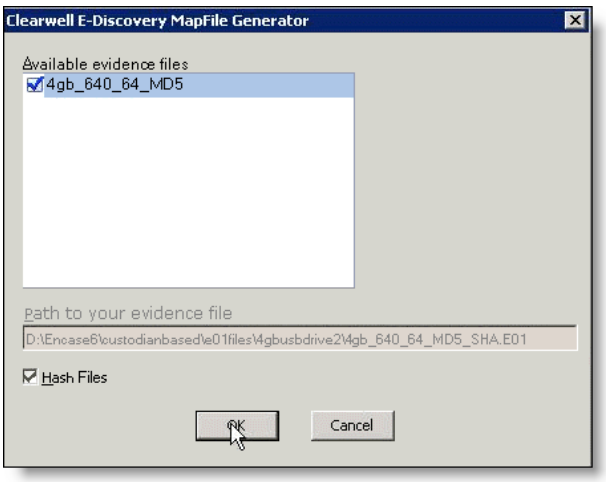
To prepare E01 files for processing

1. Take the eDiscovery Mapfile Generator setup file from ***\CW\<version>\utilities\EncaseEnscript*** and copy it to a machine that has EnCase installed.
2. Run the EnScript installer and follow the on-screen instructions.

   This installs the Mapfile generator on the machine. The Mapfile generator is, in essence, an Encase plugin.
3. Start EnCase and open the case that contains the evidence files.
4. Locate **E-Discovery Mapfile Generator** in the **EnScript** tab of your EnCase application.



5. Right-click **E-Discovery Mapfile Generator** and choose **Run** to open the Mapfile Generator dialog box.



Note: The **Hash Files** option must always be selected.

Note: **Selecting Evidence Files.** It is recommended to always hash the file first, otherwise it will be necessary to do this at the time of discovery to support the de-NIST of files which could result in slower performance.

6. Select the evidence files, select the **Hash Files** option, and click **OK** to create the MDM file.

Note: The MDM file must reside in the same folder as its associated evidence files (E01 files). As long as this is the case, the product will automatically recognize the evidence files when processing the case folder.

7. From the top navigation bar, for the selected case, click **Processing > Sources and Pre-Processing**, and add the case folder containing the evidence files and the corresponding MDM files.

   Your case folder can contain any combination of loose files, emails, email container files, and L01/E01 files. For more information, see *"Adding Case Folder Sources" on page 53*.

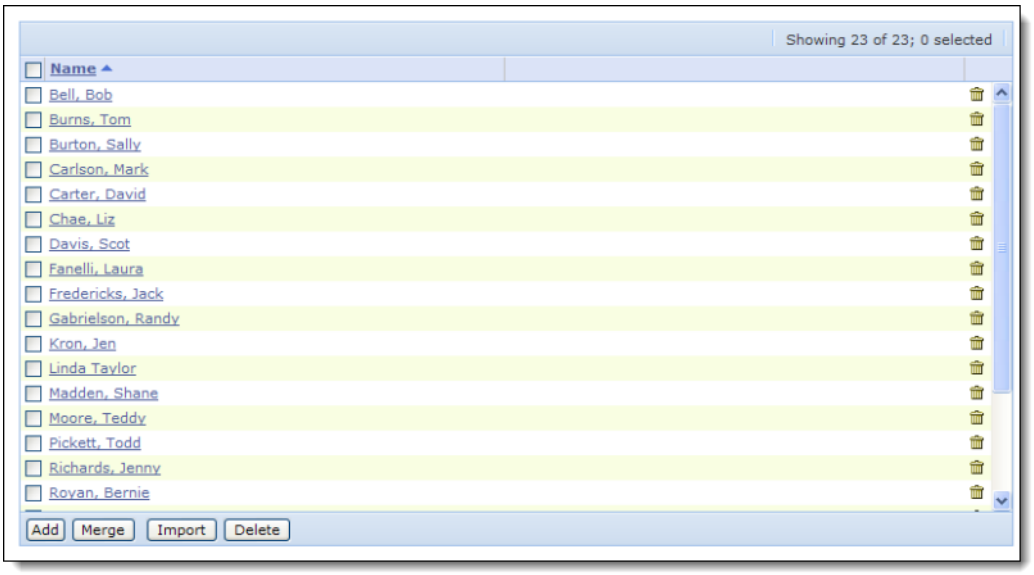
   Note: The E01 /MDM file pairs created by the MapFile Generator are portable. However, be sure to note the timezone in which the data was collected and stored in the EO1 files. The timezone needs to be set within the product to ensure the dates associated with the loose files match the information in Encase.

Defining Case Custodians

For each case, you may assign one custodian to each mailbox, file, and directory of loose files (though it is not necessary to assign a custodian to every mailbox, or file, for example). For those assigned however, you can search the case for documents associated with its assigned custodian. You can also merge custodian records of two or more custodians found to be duplicates. (See *"Merging Custodians" on page 63*.)

To define custodians for a case

1. On the navigation bar, for a selected case, click **Processing > Custodians**.



2. To add a new custodian to the case click **Add**, enter the custodian name, and click **Save**.
3. To delete a custodian, click trash icon for the name and then click **Delete Custodian.** The custodian is removed from all mailboxes, mail files, and directories.
4. To import custodians, from a file:
   A. Click **Import** to open the Import File dialog box.
   B. Choose whether to import from a text file (.txt) or CSV (.csv) file. The CSV file option allows you to include tabular data. To see a sample CSV file format, click **Download example CSV file.** The basic format for both options is one name per line/row.
   C. Click browse ... icon to select the file to upload.
   D. Click **Next** to upload the selected file. The uploaded items are displayed.
   E. Click **Finish**. The custodians are added to the custodian list on the Custodians screen.

Merging Custodians

When you have two or more of the same, or similarly-named custodians (representing the same custodian or individual) you can merge them into one unique custodian assignment. This is especially useful if you collected custodian data through the Identification and Collection module, where multiple name variations might appear for a single custodian. In this case, you may find multiple name variations on a single custodian. (Refer the *Identification and Collection Guide*.)

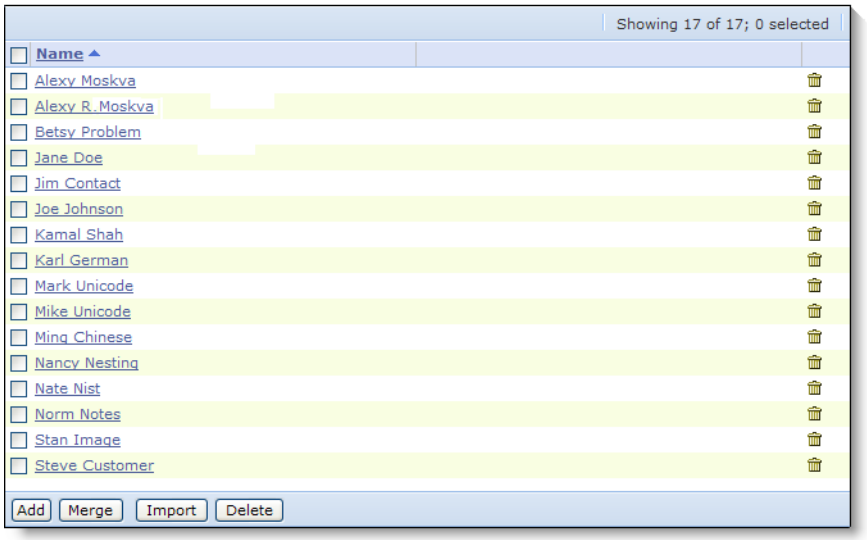
Note: Unlike custodians in your case data, custodians in the Identification and Collection (IC) module are case sensitive. As a result, custodians in IC may be merged with similar custodian names when added to your case. For example, the IC custodians "joe admin", "Joe admin", and "Joe Admin", who are all considered unique in IC, are treated as the same custodian if added to your case. Thus, if you add a collection set (created in IC) containing the custodian "joe admin" to a case that contains another custodian "Joe Admin" they are merged as one custodian. However, if that same case contains no similarly-named custodians, and all three IC custodians are added to the case, they are considered unique.

Before you begin: Merging custodians is optional, and can be done either before or after processing your case data. However, if you process your case data first, before merging custodians, you must rerun post-processing for the merged custodian assignments to take effect.

To merge custodians

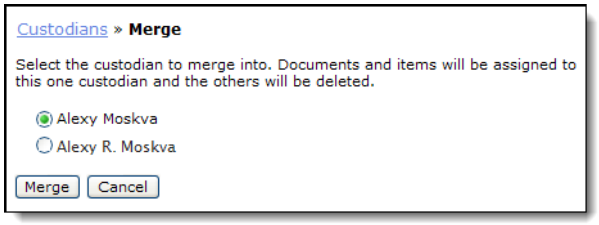
1. On the top navigation bar, for a selected case, click **Processing > Custodians**.

   An alphabetic list of custodians displays.



2. Select two or more custodians that you want to merge into the same custodian assignment.

3. Click **Merge**.



Note that the single custodian you select will automatically be associated with all documents and items previously associated with both. All other related custodians listed on the **Custodians > Merge** screen will be deleted.

4. Click **Merge.**

Note: You must run post-processing for this change to take effect. Once you run post-processing, you will not be able to undo this merge.

5. At the prompt, click **OK** to confirm the single custodian assignment.

The custodians you selected to be merged now appear in the list with the note "will merge to [new custodian name]". The merge will occur once the case has completed post-processing, but cannot be undone after post-processing. To undo the merge, see Unmerging Custodians.

CAUTION: Once merged custodians have been post-processed, documents or mailboxes associated with the original custodian records can no longer be tracked, and no historical data is retained at a file level. All documents and mailboxes will be associated with the new custodian.

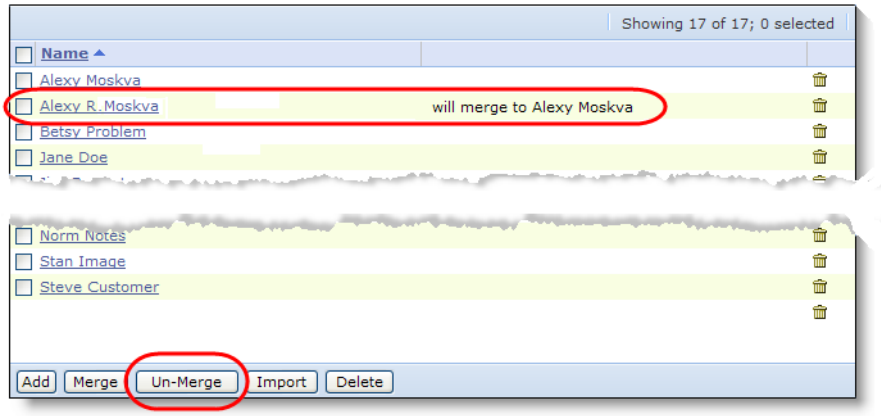
Unmerging Custodians

If you want to undo a custodian merge and have not yet run your case through post-processing, you can still un-merge the custodian assignment. In this case, the "Un-Merge" button appears at the bottom of the Custodians screen.

To un-merge custodians

1. On the navigation bar, for a selected case, click **Processing > Custodians**.

   An alphabetic list of custodians displays, indicating the custodians slated to be merged.



2. Select only the custodian(s) you no longer want to be merged (not the custodian slated to be merged to).

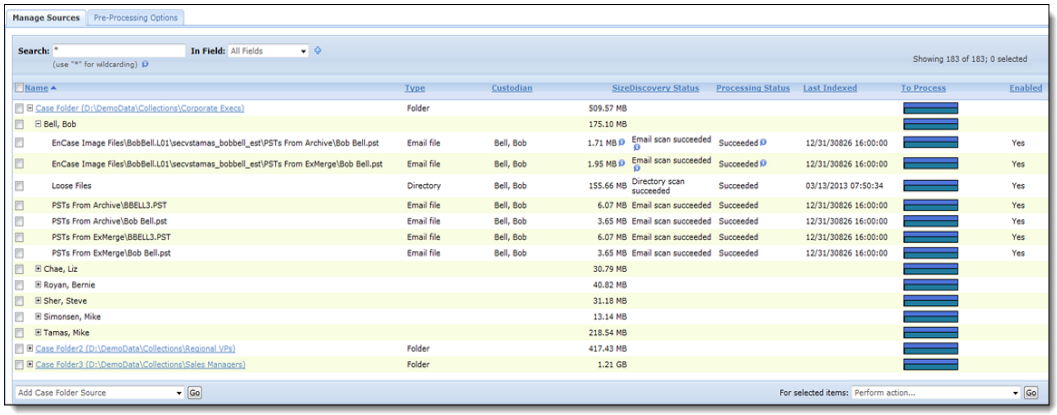
3. Click **Un-Merge**.

   The custodian merge is reversed and will not be merged upon post-processing.

Assigning Custodians

To assign custodians to specific mailboxes, mail files, or directories

1. Click **Processing > Sources and Pre-Processing** to open the Manage Sources screen.



   A. Select the check box next to the items where you want to assign the same custodian (or remove the custodian). Click the first check box to select all the items on the screen.

      To search the list:

      › From the in menu, select the column to be searched.

      › Enter the first few characters of the search text in the Search for field (use a "*" to indicate any text).

   B. Select a custodian from the menu in the lower-right corner, and click **Go**. To remove the current custodian, select **<none>**. To define a new custodian, select **New custodian**, enter a custodian name, and click **OK**.

   C. Click **Manage Sources** to return to the list of document sources.

2. To apply the custodian changes to the entire case, click **Rerun Post-Processing**.

   Note: In version 7.0 and higher, you can search for custodians from within the Manage Sources tab, if for example, you had a collection set which contained multiple loose files, each with its own custodian and wanted to see which files belonged to specific custodians. For more information about collection sets, refer to *"Creating, Analyzing and Processing Collections" in the Identification and Collection Guide*.

Pre-Process Your Source Data

Note: To view and use the **Pre-Processing Options** tab, the Pre-processing module must be installed, licensed, and enabled on your system. Settings on the tab are disabled if a discovery or processing job is currently running for the case.

How Pre-Processing Works

When you create a source for a case, the system determines the data set size and number of files. The results of these calculations are shown visually as bars on the **Processing Options** screen (from left to right in the **Summary** view):

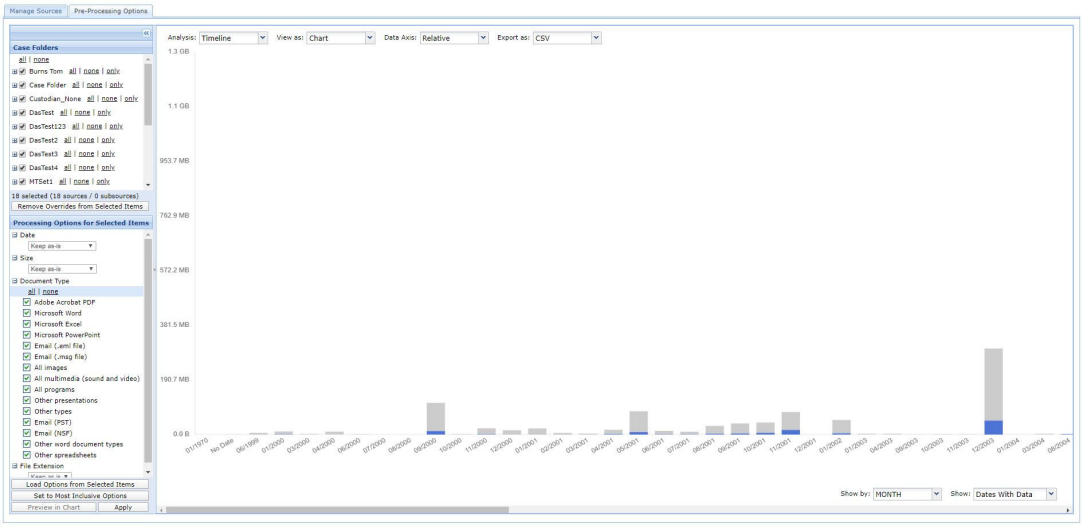
- **Source Files** — Total size of the original data set (number of files and size).
- **Change Due to Extraction** — Change in the data set size due to file extraction.

  For example:

  - If the data set contains zip files, extraction will cause the total file count and size to increase.
  - If the data set contains PST, NSF or container files, the count and size of the data set after extraction will be less than the count and volume of items in the original source location. The compression of the data set accounts for this difference and, depending on the compression ratio can substantially decrease the size of the original and uncompressed files.
- **Excluded Known Files —** Change in the data set size due to the exclusion of known files from the NIST list and any other added file hashes.
- **Preprocessing Errors**— Change in the size count rejected due to errors during preprocessing. Includes corrupt files, password protected files, and unrecognized files.
- **Already Processed** — Change in the data set size due to the exclusion of files that have already been processed.
- **Excluded by Processing Options** — Change in the data set size due to the exclusion of document types on the **Add Case Folder Source** screen or based on the processing options that are set for subsources that are accessible from the **Manage Sources** tab.

- **To Process** — Size of the remaining data set that will be processed.

  You can use pre-processing to examine statistics prior to processing and to examine what has been processed after the processing operation is complete. In the following figure, the unprocessed portion of the data set is shown in gray.



Setting Up Pre-Processing

You can set up processing criteria to identify files and/or document types that you want to exclude from processing.

You can explicitly include or exclude documents from processing in either or both of the following ways:

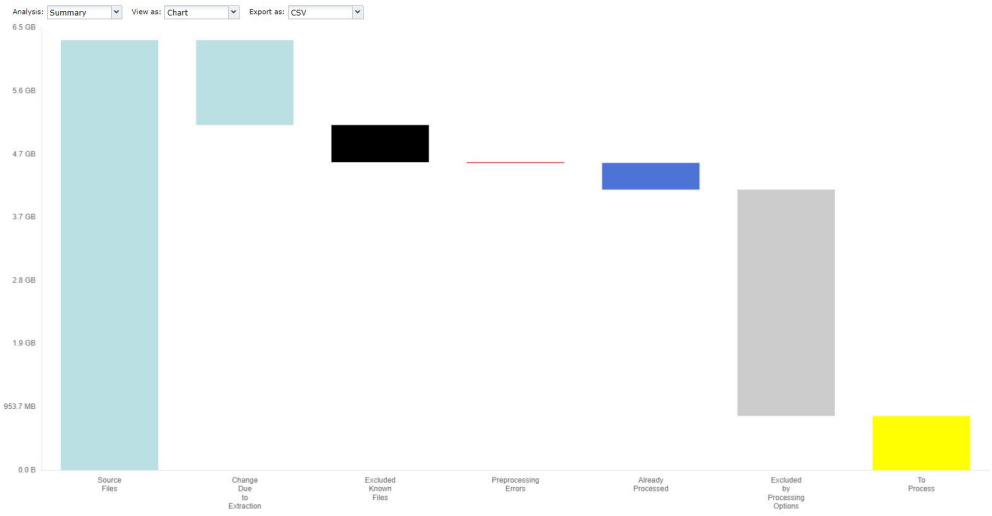
- Specify the file types under **Processing > Sources and Pre-Processing > Manage Sources > Add Case Folder Source** screen. You can identify specific document types, such as email files (.eml/.msg) or Adobe Acrobat PDF files, to exclude from processing. See *"Pre-Processing Options Tab" in the next section*.
- By specifying known files under **System > Known Files**. By default, the known files list includes the NIST list, a set of md5 hashes from the National Software Reference Library (NSRL) of the National Institute of Standards and Technology (NIST). These hashes are for common programs, such as Microsoft applications, that are generally not required as part of an eDiscovery search. By including the NIST list, you can reduce processing times without jeopardizing the integrity of your searches. If you have additional sets of files to exclude from processing, such as a corporate laptop image, you can create a hash of the image, add it to a CSV file, and then add the CSV file as another known file. Refer to *"Defining System Settings" in the -System Administration Guide*.

The query is ANDed across options and ORed for values within a setting. For example, to include only pdf extensions, select the document type “adobe pdf” and choose only pdf for the extension.

Pre-Processing Options Tab

The **Pre-Processing Options** tab next to **Managing Sources** presents the current processing settings in the following views:

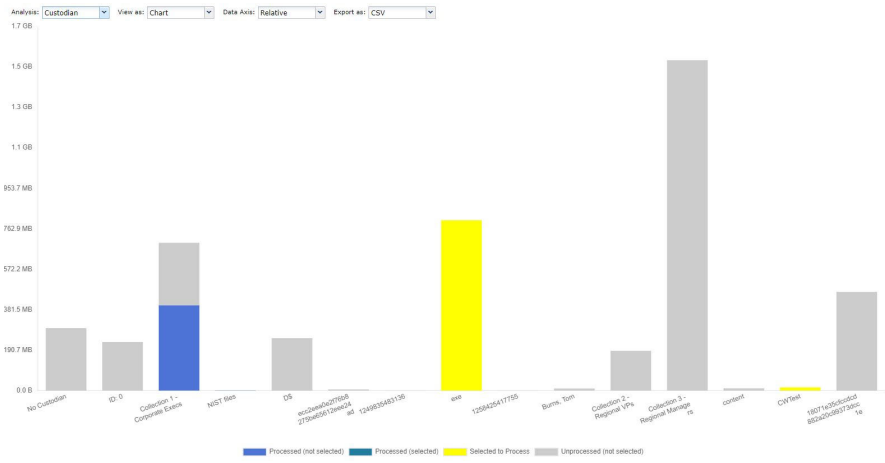
- **Summary**—Information based on total data set sizes. This is the view that is displayed when you first open the **Pre-Processing Options** tab.



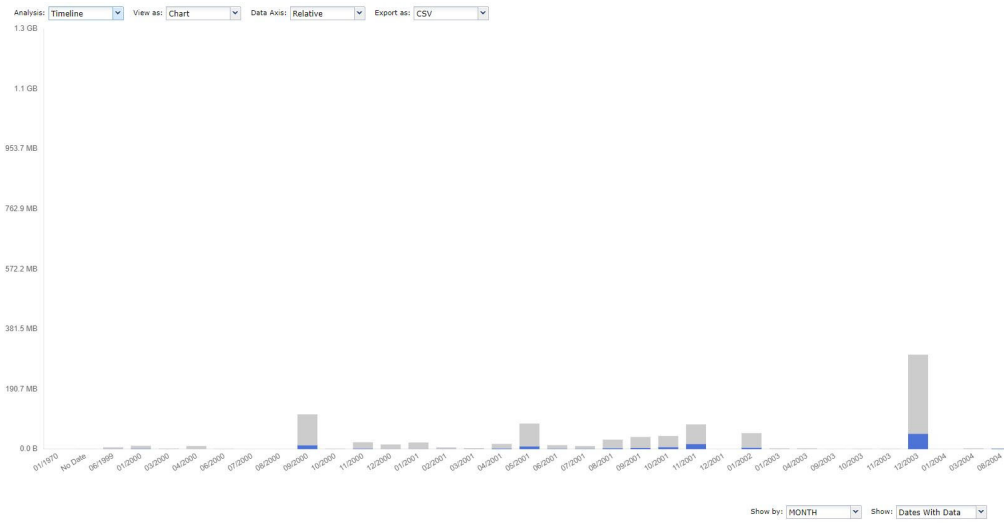
If errors occur during pre-processing, a red bar appears indicating the total number of errored files. This can include any corrupt, password-protected, or unrecognized files.

- **Document Type** — Information presented by file types. The file types are determined by actual file signature, not by file extension.

- **Custodian** — Information presented by custodian name.



- **Timeline** — Information presented chronologically.



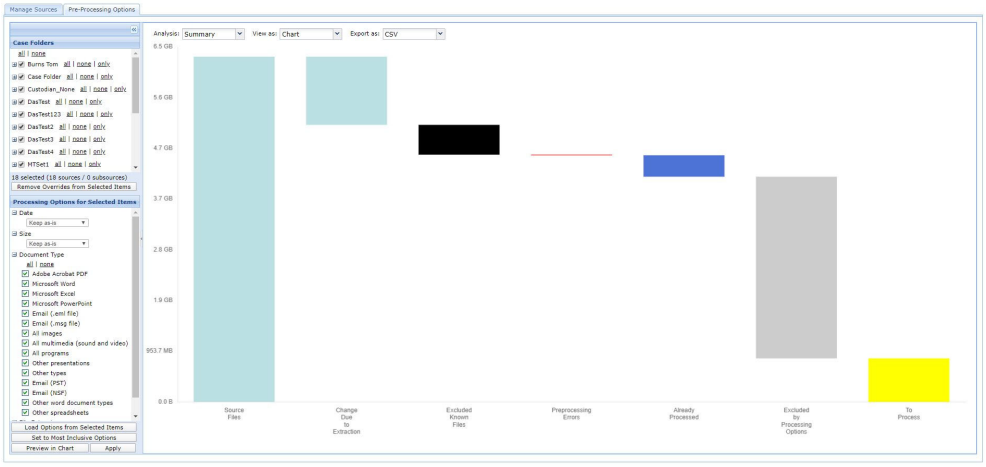
Setting Pre-Processing Options

Use the **Pre-Processing Options** screen to perform the following functions:

- Perform analysis of documents to identify potential issues with the document set and estimate processing costs.
- Set processing options to restrict the set of documents that are processed into the product.

To use the Pre-Processing Options tab

1. On the navigation bar, for a selected case, click **Processing > Sources and Pre-Processing**.
2. Click the **Pre-Processing Options** tab.



3. To assist in document analysis, select viewing options using the controls at the top of the tab (and bottom of the tab for the Timeline view), as described in the following table.

Viewing Options

Options	Description
Analysis	Choose a view for the data (summary, document type, custodian, or timeline), as described at the beginning of this section.
View As	View the data in a chart or table.
Data Axis	Base the units of the vertical axis on the overall relative, or case size to the largest data point.
Export	Export data in a CSV or XLS file. When you choose one of these options, a pop-up window opens with a prompt to open or save the file.
Show by	(Timeline view only) Show units in days, months, quarters, or years.
Show	(Timeline view only) Display all dates, or only dates that have data.

4. To select case folders and items for pre-processing, refer to the following table.

Case Folder Selections

Options	Description
Case Folders	Choose the folders that you want to include for pre-processing in any of the following ways: • Select check boxes for the folders. To expand a folder and show the subfolders, click the + sign to the left of the entry. Click the - sign to collapse the entry. • Click **all** to select all of the files in a given case folder or subfolder. • Click **none** to exclude all of the files in a given folder or subfolder. • Click **only** to limit the files to only those that are explicitly selected. • Click by **Custodian** to open a pop-up window that contains the list of custodians. To add custodians, first click **none** to clear all of the folder selections, Then with the pop-up window open, click **Include** to explicitly include the files associated with that custodian. As you make selections, the chart (or table) is updated automatically. Click **Exclude** to remove the files for that custodian. • Click **Discard Overrides on Selected Items** to remove custom selections on the selected items.

5. To set processing options for the selected items, refer to the following table.

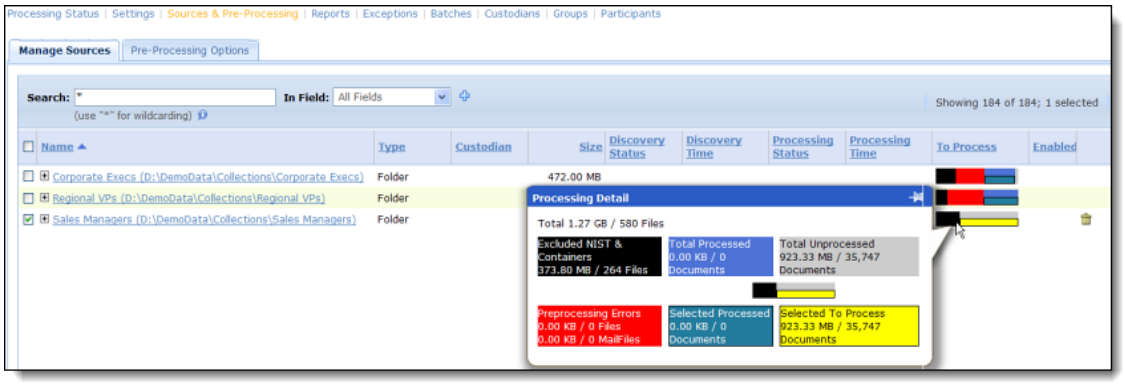
Pre-Processing Options

Options	Description
Pre-Processing Options for Selected Items	Choose the Pre-processing options for the folders that are selected in the Case Folders area. **Note:** The date that is used is the modified date of the file in the case of loose files (including loose file email messages), or the sent date of the email if it is in an email container (PST or NSF file). • **Date**—Choose one of the following date options: – **Keep as-is**—Use the date specification that is already in effect for the selected subsources, as opposed to overriding them with values that you enter on this screen. – **All dates**—Use all dates in the data set. – **Dates On or Before**—Include all dates on or before a specified date. To select a date, use mm/dd/yyyy format or click the calendar icon. – **Dates On or After**—Include all dates on or after a specified date. To select a date, use mm/dd/yyyy format or click the calendar icon. – **Dates Between**—Include all dates between specified dates. To select a date, use mm/dd/yyyy format or click the calendar icon. • **Size**—Choose one of the following options: – **Keep as-is**—Do not change the size configuration. – **All Sizes**—Use all file sizes in the data set. – **Sizes Smaller Than**—Use all file sizes smaller than the selected size. – **Sizes Larger Than**—Use all file sizes larger than the selected size. – **Sizes Between**—Use all file sizes between the selected sizes. • **Document Type**—Select check boxes for the document types that you want to include in processing. To select or deselect all of the document types, check or clear the check box at the top of the list. Some check boxes support three states: checked, unchecked, and grayed to indicate that the value is kept as-is. For three-state check boxes, a drop-down menu is shown when you click the check box to allow you to choose among the three states.
	Note: The document type applies only to the top-level document type of a loose file or message, and does not apply to attachments. If you choose to process a particular type of email, the email messages and all of their attachments will be processed, regardless of whether the document type of the attachment is excluded here.
	• **File Extension**—Select Keep as-is, Exclude or Include.
Populate Values from Selected Items	Use all the configured settings to determine the data set to be processed.
Set to Most Inclusive Values	Apply the most complete configured values across all of the selected folders.

Pre-Processing Options (Continued)

Options	Description
Preview in Chart	Refresh the chart to reflect all of the configured settings. Items shown in gray will not be included in processing. If **Set to Most Inclusive Values** is clicked, then all relevant settings are included, and none of data is shown in gray.
Apply	Save the configured settings to use in processing.

You can view a summary of the pre-processing results by moving your cursor over an entry in the **To Process** column on the **Manage Sources** tab (**Processing > Sources and Pre-Processing > Manage Sources**).

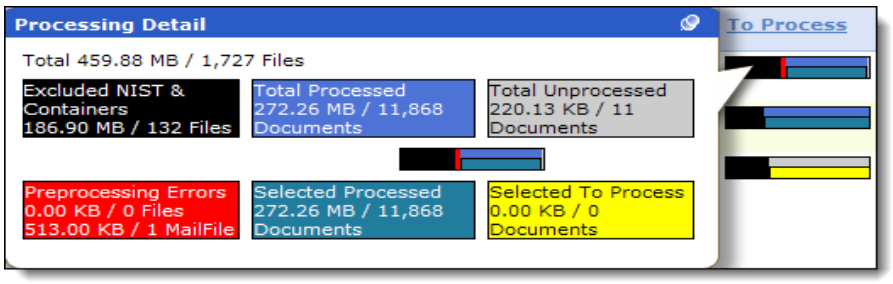


Viewing Processing Detail

The bar in the **To Process** column is also shown in the information bubble that is displayed when you move your cursor over an entry in the column at any level of the case folder hierarchy. The image is divided into the following sections:

- **Excluded NIST and Containers** (black): Known NIST documents and containers excluded from processing as a proportion of the total set of files.
- **Total Processed** (blue): Documents that have already been processed as a proportion of the total set of files.
- **Total Unprocessed** (gray): Documents that have not been processed as a proportion of the total set of files.
- **Selected Processed** (blue/green): Documents that have already been processed as a proportion of the total set of files selected for processing.
- **Selected to Process** (yellow): Documents that have been selected for processing but not yet processed as a proportion of the total set of files selected for processing.
- **Preprocessing Errors** (red): Documents that caused errors due to corruption, password protection, or unrecognized types or formats, for example. If errors occurred during pre-processing, this view provides an opportunity to attempt to correct the errored files before processing.

  Note: Errors that do not have an error code or file ID are not included in the error count, and will not appear in the processing detail view.



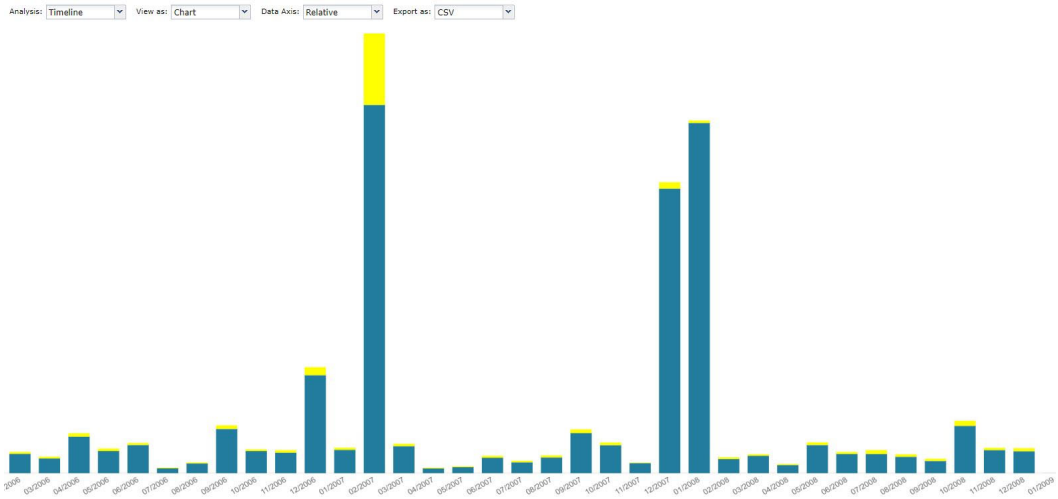
Pre-Processing Example

The example in this section shows some of the ways that processing options can be used.

Assume that managers have added additional documents (a second collection) into an original case. Because the collection process was prioritized, documents for the first collection have already been received and processed.

For the new collection, documents are required for 2006 to 2008; however, the new collection includes documents for 2005 to 2008.

The user excludes the documents for 2006 by specifying **Dates On or After 01/01/2006** in the Processing Options area (see *"Pre-Process Your Source Data" on page 67*). This yields the following pre-processing timeline view, which shows the excluded period in gray.

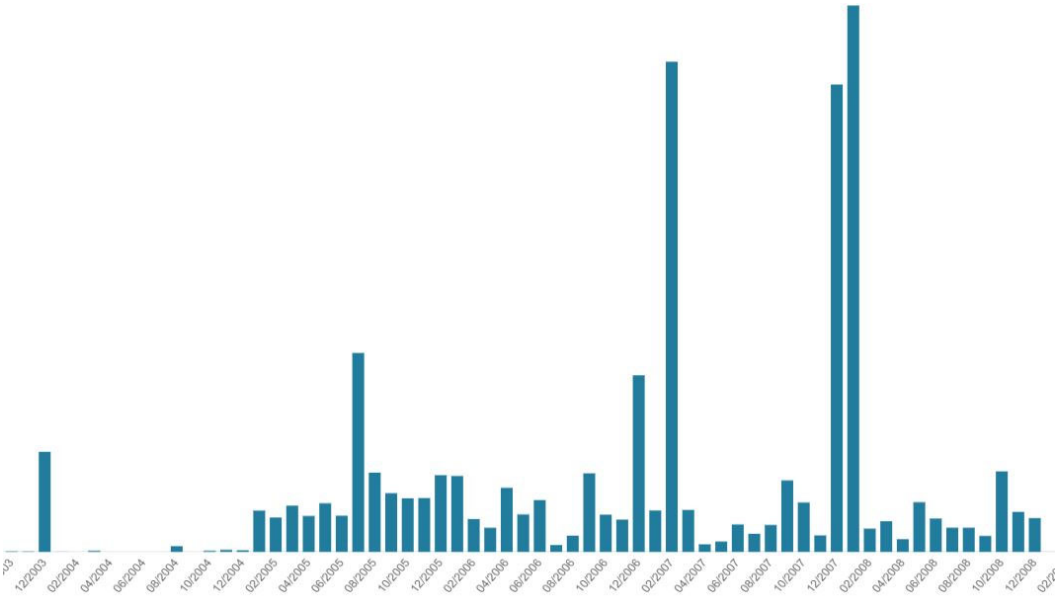


Pre-Processing Options Examples

The summary chart for the new collection shows the exclusions (de-NIST and processing options). Pre-processing yielded approximately a 61% volume reduction.

Finally, the timeline view for one of the custodians shows four months of missing documents in the collection.

The pre-processing analysis charts can be used to identify potential issues in the collected data. For instance, in this example a customer is missing one month (06/2004) of email messages because of a missing PST file in the collection. Because the pre-processing analysis goes down to the individual email level, you can quickly and easily identify such gaps.



Information Classification

Note: If you are not familiar with basic classification operations and terminology, see *Veritas Information Classifier Help*.

Overview

Starting with version 9.0, the eDiscovery Platform lets you automatically classify sensitive and critical case data based on a set of built-in and custom policies.

Note: The Information Classification feature can only be enabled for cases created in version 9.0 and later. Cases created before version 9.0 are not supported.

The platform integrates with Veritas Information Classifier (VIC) to analyze and classify eDiscovery data. VIC uses both predefined and user-defined policies to assign classification tags to your eDiscovery data during the processing phase. Once these tags have been applied, users can view pre-selected classification filters (system tags) in the Analysis and Review mode to quickly identify documents that match the VIC tags.

Note: Classification filters operate like predefined system tags in the eDiscovery platform.

Considerations

- This feature is only available for cases created in version 9.0 and later. Information Classification does not support cases that were created before version 9.0.
- In a Distributed Architecture configuration, the Information Classifier configuration is accessed on the cluster primary node only.
- Once policies are enabled, the policies will be used for ALL VIC-enabled cases on the system.
- If OCR is disabled during processing, and the OCR is used via **Actions >OCR** within the case, any OCR data will also have the classification policies applied at that point.
- Classification will have an impact on processing performance. The more policies that are selected, the more work the engine needs to do adding overhead to processing time.To limit the impact, only enable policies you know you will need to use.
- Once a case has been setup for classification, and the first batch of case data has been processed, the policies that were enabled at that point in VIC will stay with the case. This means that any future enabling of policies in VIC will have no effect on future data ingested into that case.

  New case behavior is different in that any new cases will be associated with the policies enabled at that particular point in time.
- For a case where classification was enabled and data ingested by enabling a certain set of policies, if one or more policies are disabled in VIC, future data ingested into that case will not make use of those disabled policies for classification.

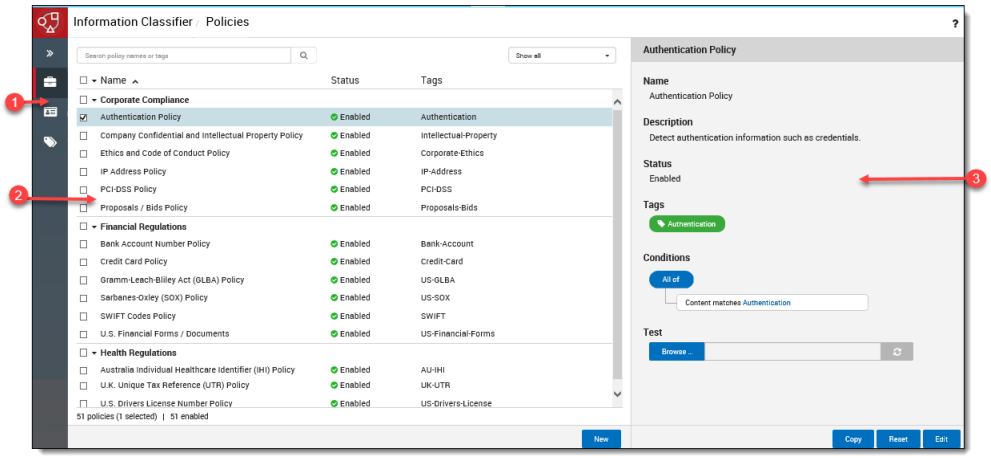
Setting Up Veritas Information Classification Policies

Classification policies are all managed by the Veritas Information Classifier (VIC) interface. VIC is a common interface that is shared by a suite of Veritas applications including Enterprise Vault and Data Insight.

To open the Veritas Information Classifier

For security reasons, you cannot access this interface remotely. You must be logged on to the eDiscovery desktop server.

1. From the eDiscovery Platform server desktop, launch the browser
2. In the address bar, enter the VIC URL: http://localhost:8090/vic/#!/admin/policies

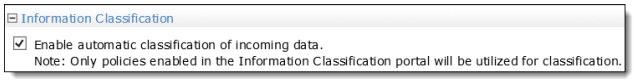


The VIC window has three areas:

1. **Navigation bar:** The navigation bar provides buttons with which you can open the Veritas Information Classifier pages. You can collapse the bar so that only the buttons show, or pin the bar so that it remains expanded while you work.
2. **Policy list:** The item list provides a list of the available items, together with basic information about them. Click an item to view more information in the details pane. The controls at the top of each list let you search for items by name, filter the items according to various criteria, expand and collapse the list, and change the sort order.
3. **Policy details pane:** Provides extensive information on the selected policy. You can use this pane to edit an item or create a new one to add to the list.

Get Started with Information Classification Workflow

Step	Action
Step 1	Navigate to the Veritas Information Classifer (VIC) Policy Manager to access built-in and custom Information policies.
Step 2	Enable/Disable Policies that meet your organization's needs. All policies are disabled by default. **Note:** To avoid impacting performance, be judicious in your policy selections.

Step	Action
Step 3	In the eDiscovery platform, enable Information Classification in the case settings as part of case processing setup. By default, Information Classification is disabled.  For eDiscovery platform settings, see *"Information Classification" on page 24*
Step 4	Proceed with normal case processing.

Enabling or Disabling Policies

Initially, all the policies are disabled. You must enable a policy if you want the Veritas Information Classifier to check for and tag the items that match the policy.

Note: Enabling a lot of policies can affect performance. In addition, policies with complex conditions take longer to process than those with simple conditions

To enable or disable a policy

1. At the left of the Veritas Information Classifier, click **Policies**.
2. Select one or more policies that you want to enable or disable and then click **Edit**.
3. You can enable or disable multiple policies at once.
4. In the Status field, select **Enabled** or **Disabled**.
5. Click **Save**.

Image Overlay

The Image Overlay action allows you to easily replace the native images accessible in review mode with one or more images from an external tool. The integration of image remediation with review means that, in addition to importing bulk and multiple images from the **Processing > Imaging and Rendering** screen, you can also import an image on an item-by-item basis while in review mode.

Before You Start

The Image Overlay manages images from a central and shared location that you create and provide. This location needs to be both available and reliable. For cluster environments, this should be a shared location. To access images in cluster or single appliance environments, you will want to set the image remediation system property. Make sure you create a shared directory location for the native image files before you set the property.

Note: For cluster environments the entry must be a shared location.

Note: You must have the **System Manager** role for the following:

To set image remediation property for image overlay:

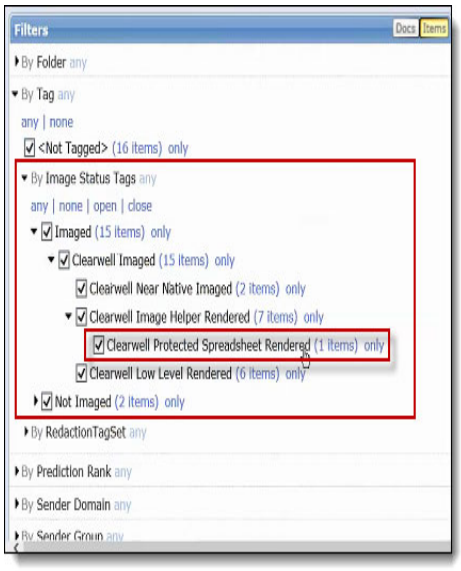
1. From **System > Support Features**, select **Property Browser**.
2. Enter the property:
   `esa.image.remediation.individual.import.shared.location`
3. For the value, enter a shared location (cluster) or a local shared directory.
4. Click **Submit** to save your setting.

Replacing Native Images

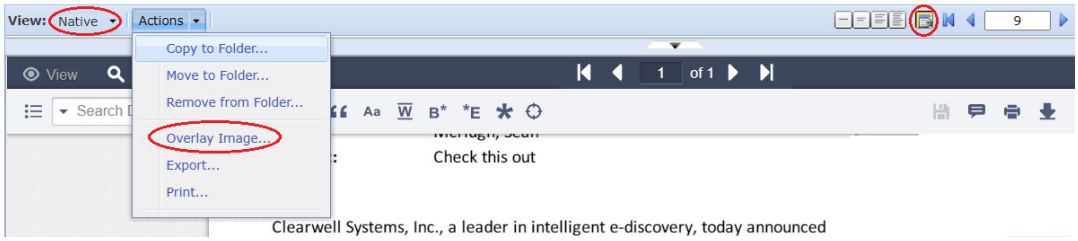
To Overlay an Image

1. From the **Analysis and Review** screen, select the item from the Filter Search Results that you wish to replace. In addition to your own tagging selections, there are several system generated image tags (under the heading Image Status Tags) that identify the status of images and can assist in locating the item you wish to overlay. For a description of all the system image tags, see *"Image Remediation" on page 88*.

   Note: You can disable default loading for Find Similar. Go to **Processing > Settings,** under Find Similar Settings, and check the Disable find similar in Review mode. Please be aware that if Find Similar is disabled, it is no longer visible in either the header section and Related Items pane.



2. Once the item is in Review mode, select **Native** for the view.



3. From the Actions menus, select **Overlay Image...**

   Note: You can overlay a file with multiple external image files. The imaged files are numbered per page and follows the page order on the Overlay Image screen.

   The Overlay Image screen displays and requests a file to upload.

4. Browse to the externally imaged file that you want to use to replace the current native file image and click **Next**.

   An Overlay Image Warning screen displays to confirm your choice to replace the image.

5. Click **Yes** to replace the image or **No** to discard your selection and start over.

   Note: You cannot overlay images for redacted images or images from a production folder If you do so, the operation will fail.

   After clicking **Yes**, the overlay image replaces the current image. If the image overlay was unsuccessful then an error message displays with an explanation for the failure. For example: `Image import error: Image(s) cannot be imported as it has been produced.`

Image Remediation

Image Remediation integrates all the actions that manage the workflow for handling native images. This workflow integration means that users can easily move native images through the various stages of identification, export, import, delete and reporting using a central command and control menu.

There are many advantages to this approach. For example, there may be situations where using an external tool to image documents is preferable to using the internal TIFF image handling capabilities of the product. Candidates for external imaging might fall into the following categories:

- An unsupported item type (such as Microsoft OneNote)
- An item type that is not imageable (for example, DLLs, EXEs, audio, or video files)
- Special image quality requirements that exceed the product's production capabilities (such as the exclusion of blank pages or setting print parameters that are different than Excel spreadsheet defaults)

In these situations, image remediation provides a workflow that easily identifies the status of items with system image generated tags. Once items with imaging issues are identified, you can select them for external TIFF imaging and export the items, in native form, to the TIFF imaging tool of your choice. You have the ability to export just the items you need instead of the entire document family containing those items. Once you have externally generated the TIFF image files, you can import the images back into the product individually or in bulk fashion. For your convenience, the image remediation feature tracks this information online and also provides a report of all internally and externally imaged items.

You can manage all of the image related tasks from the **Imaging and Rendering** screen. You can take the following actions:

- *"Generating an image status report" on page 92*
- *"Native Image Caching For Faster Review" on page 92*
- *"Deleting Native Images" on page 93*
- *"Exporting Native Images for External Imaging" on page 94*
- *"Import Native Images" on page 94*

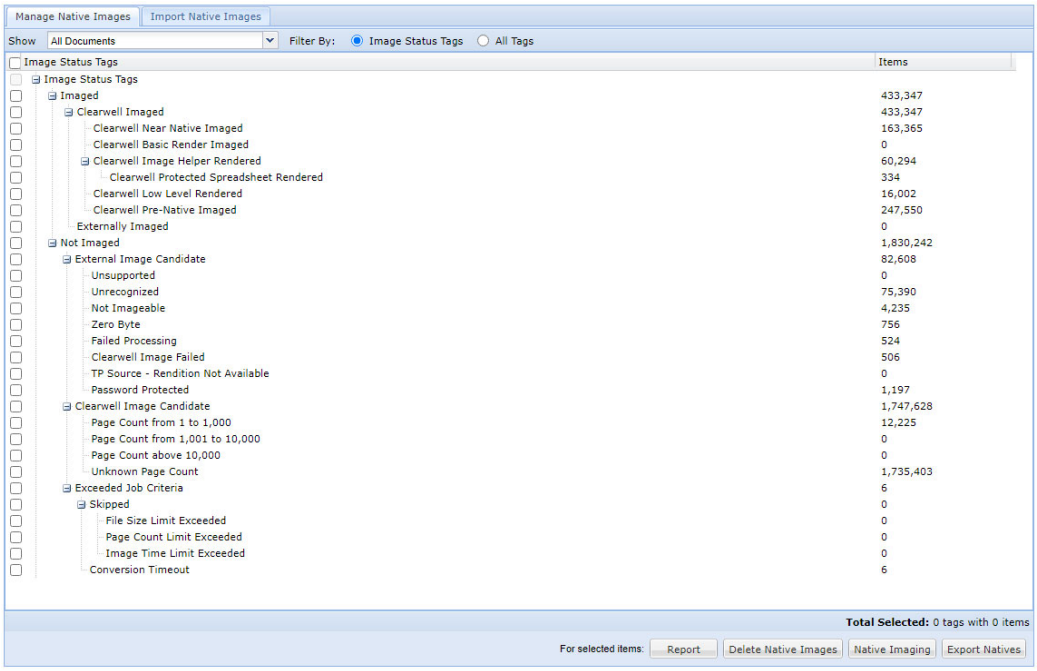
Note: If the case is created in pre-10.1 release, then after upgrade to release 10.1, the following operations are not available until Imaging Tool Upgrade is run for that case:

- **Manage Native Images** > **Delete Native Images**
- **Native Imaging** > **Manage Native Images**
- **Import Native Images**

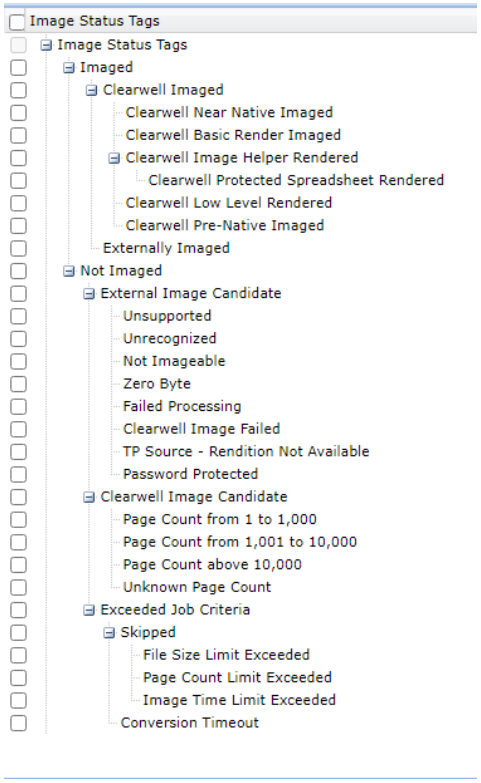
See the *Imaging Tool Upgrade Guide* for details.

Manage Native Images

The **Manage Native Images** screen (**Processing> Imaging and Rendering**) enables you to perform all of the above actions except the ability to perform imports. You can perform imports from the **Import Native Images** screen.



The Image Status Tags provide helpful image filtering.



These system-generated image status tags are described in the following table.

Manage Native Images Tags

Status	Description
Imaged	Items successfully imaged.
Clearwell Imaged	Items for which images were *not* imported through "image overlay" workflow and where native image generation was successful.
– Clearwell Near Native Imaged	Items whose images are successfully rendered with a near native quality. See *"Manage Native Image Tag Considerations:" on page 92* for clarification on items from restored cases or system upgrades.
– Clearwell Basic Render Imaged	Items whose images encountered issues and failed initially, but were successfully rendered with a text or rich text quality after retry attempts.
– Clearwell Image Helper Rendered	Indicates item whose images were rendered in rich PDF format.
– Clearwell Protected Spreadsheet Rendered	Indicates Excel protected spreadsheets whose images were rendered in rich PDF format, but the imaged PDF excludes any hidden rows and columns if they are present in the source file.

Manage Native Images Tags (Continued)

Status	Description
– Clearwell Low Level Rendered	Indicates all items whose images were rendered in low level PDF format.
– Clearwell Pre-Native Imaged	Indicates all items whose images were rendered in rich PDF format generated by PrizmDoc.
Externally Imaged	Items whose images were imported through the "image overlay" workflow.
Not Imaged	Items which do not have images (neither generated internally or externally imported).
External Image Candidates	Items that cannot be imaged by the eDiscovery platform.
– Unrecognized	Item types not recognized by Veritas eDiscovery Platform.
– Unsupported	Items types that are not supported (such as Microsoft OneNote).
– Zero Byte	Items with zero-bytes.
– Not Imageable	Item type not imageable (such as DLLs, EXEs, audio or video files).
– Clearwell Image Failed	Items that could not be imaged for reasons other than those listed above (for example, due to some type of product issue).
– Failed Processing	Items that failed processing (for example, password protected attachments).
– TP Source - Rendition Not Available	Load File Import (LFI) source where neither Image/Native/Metadata rendition is available.
Clearwell Image Candidate	Items that can be imaged in Veritas eDiscovery Platform.
– Page Count 1 to 1,000 pages	By default, page counter is enabled for PowerPoint and PDFs.
– Page Count 1,001 to 10,000 pages	
– Page Count above 10,000	
– Unknown Page Count	Lists documents where the page count is unknown. For example, emails and images.
Exceeded Job Criteria	
Skipped	Items did not meet user specified limits for the imaging job(s). Items could match more than one of the user specified limits (see those mentioned below).
– File Size Limit Exceeded	File size is greater than user specified limit for the job.
– Image Time Limit Exceeded	Expected imaging time is greater than user specified for the job.
– Page Count Limit Exceeded	Page count is greater than user specified limit for the job.
Conversion Timeout	Conversion has timed out because the items took longer than expected to process.

Manage Native Image Tag Considerations:

- These tags are not displayed in the review dash board or in case reports.
- Use of page counts to classify or analyze documents during native imaging is not supported for LFI sources.

Generating an image status report

The image status report is an essential aid for image tracking and monitoring. One of the most useful aspects of the report is its ability to identify items that are good candidates for external imaging tools. The report allows you to quickly glean insights and an understanding of why a particular item or items have not been imaged. In addition to summary data, each item is assigned a status category and a specific reason as to why it was not imaged. For example, you can drill down on an item under the "Unsupported" category to find out which items and their associated file extensions are not supported. This allows you to analyze problematic images and to pinpoint actions that you need to take to resolve the issue.

To view the image status report

1. From the **Imaging and Rendering** screen, click **Report**. You must select at least one tag in order for the report to display.

   The Image Status Report displays.

2. View the report. The information is organized into the following categories:

Image Status Report

Field	Description
Item Doc ID	Assigned Item Doc ID for the item
File Name	File name for the item
File Type	Specific item file type for the item (such as PPT, DOCX, XLS, PPT, PST, PSQ, DOC)
Size	Size of the item in bytes or MB
Page Count	Count of pages
Image Status	Not Imaged
Image Sub-Status	See *"Manage Native Images Tags" on page 90*

Native Image Caching For Faster Review

Reviewers can sometimes experience extended access times for native images in part due to document retrieval performance. Instead of waiting to retrieve documents from a server or appliance, the Native Imaging Caching feature can serve native images locally to speed up delivery. This means that all subsequent viewing of native images comes from cache and not from the location where the images physically reside. This capability improves performance and gives your reviewers a much faster viewing experience.

The items for native image caching can be selected from the list of tagged items presented on the **Imaging and Rendering** main screen or by selecting the tagged images contained in a folder from the search result box. The items you select for native imaging are cached to improve retrieval and review speed.

Note: The document family of the item is submitted for caching.

Since the task of caching native images has the potential to be a resource intensive operation, the Native Imaging dialog on the **Native Image Advanced Settings** menu allows you to both manage and control various aspects of the task. In addition to specifying criteria (such as page count, file size), there are also time-dependent options. You can schedule the native image caching job to be released from the queue at a particular time, run for a certain amount of time or complete. This flexibility can help you balance the caching workload.

Note: Counts for **Total unique items selected** on the **Native Image Advanced Settings** menu can be greater than the total items selected for caching since caching is performed at a document family level.

To perform native imaging

1. From the Manage Native Image screen, select the item(s) or folders for native imaging and click **Native Imaging**.

   The Native Imaging screen displays.

2. To perform tasks on this screen, refer to *"Accelerating Review with Caching" in the User Guide*.

3. Click **Cache** to continue or **Close** to cancel.

4. If you selected Cache, an Imaging screen displays: "The Cache job has been submitted successfully", click **OK**.

   The job's completion time will vary depending on the number of items.

5. Click **Jobs** at the top of the screen to view the status of the cache job. Check the job log for errors.

Deleting Native Images

Deleting native images allows you to delete both internally and externally generated native images. There may be times when you need to delete or perhaps want to replace a cached native image or images. The **Delete Native Images** functionality provides the flexibility to perform this task from the Manage Native Images screen.

Note: Delete Native Images only deletes the cached image for both the external and internal native images. It does not delete the source of the import native image file. Additional cleanup is performed for entries made for externally imported native images.

Note: If the case is created in pre-10.1 release, then after upgrade to release 10.1, the **Manage Native Images** > **Delete Native Images** operation is not available until Imaging Tool Upgrade is run for that case.

To delete native images

1. Select the tagged native image items you want to delete from the **Manage Native Images** screen.

2. Click **Delete Native Images**.

The **Delete Native Images** menu displays with a summary detail of the Folder, Selected Tags and Items.

3. Click **OK** to confirm that you want to proceed with deleting images or **Cancel** to cancel the job and return to the **Manage Native Images** screen.
4. The **Job started menu** displays indicating that the "Delete native job started".
5. Click **Jobs** at the top of the screen to view the status and the summary and detail excel report for status.

Exporting Native Images for External Imaging

The Export Natives Images feature allows you to select those items (and not the associated document family) that require external imaging. For example, if an email has six attachments and only the third attachment requires imaging in an external tool, you can select it with item-based tag for exporting. A **Select Export Location** menu allows you to direct the export to locations that are available from the group access configuration or from external export locations set up through the property browser.

Note: The Opticon load file is exported in the necessary image format.

Document Identifiers and Image Remediation

Throughout a variety of workflows, the product generates and uses a unique identifier called a Document Identifier (Doc ID). During exports and imports of native images, the Doc ID is used for mapping and tracking purposes. The Doc ID has a numerical prefix that identifies the document family and, after the dash, a numerical suffix that identifies the item.

Here is an sample Doc ID for the third item of a document family: 0.7.13.21424-000003.

To export native images

1. In the **Manage Native Images** view, select the items you want to export.
2. Click **Export Natives,** choose the location from the **Select Export Location** menu and click **Select**.

   The **Job started menu** displays informing you that an "Export native Images job started". This starts a Metadata export job.
3. Click **OK** to confirm. You have now exported the native images. The exported items are now in an export (ZIP) file.
4. Click **Jobs** at the top of the screen to view the export file (ZIP). You have the choice to open or save it. If the exported content exceeds 2 MB, then the export file can be accessed directly from the appliance.

   Note: To access the export file directly from the appliance, click the infobubble near the export file for the location.

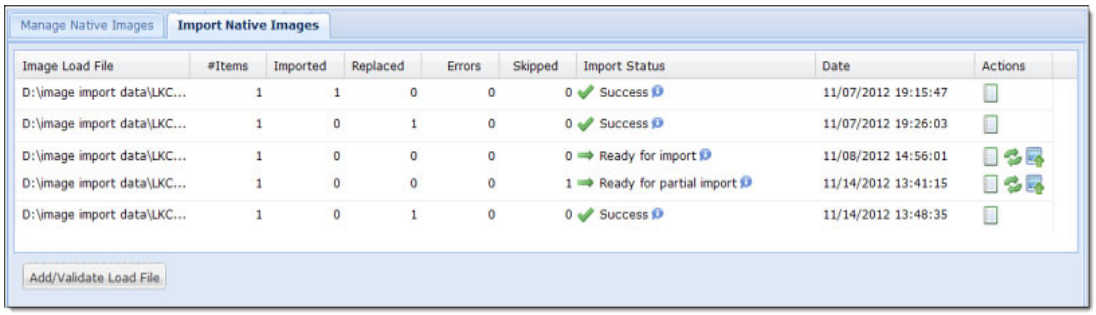
Import Native Images

Documents and items that are imaged with external tools can be imported back into the application in the following ways:

- Bulk import with a load file
- Individual import while in Review mode. For more information on this workflow, see *"Redacting Items" in the User Guide*.

Note: If the case is created in pre-10.1 release, then after upgrade to release 10.1, the Import Native Images operation is not available until Imaging Tool Upgrade is run for that case.

The **Import Native Images** screen enables you to perform bulk import from a Opticon format load file.



Once you select **Add/Validate Load File**, the import load file process goes through a series of steps. At a high level, these steps are:

1. Validation of load file
2. Import of native images load file
3. Create import status report

Load File Format Guidelines

- Image load files are in Opticon format (DAT, OPT file)
- Opticon load files can be either multi-page or single-page TIFF files

This table provides a description of each of the columns of the Opticon file used during image import.

Note: A more extensive set of Option file data fields (than the ones used for image import), are used for load file imports. For more information, refer to *"Opticon File" in the Load File Import Guide*.

Opticon Data Fields for Image Import

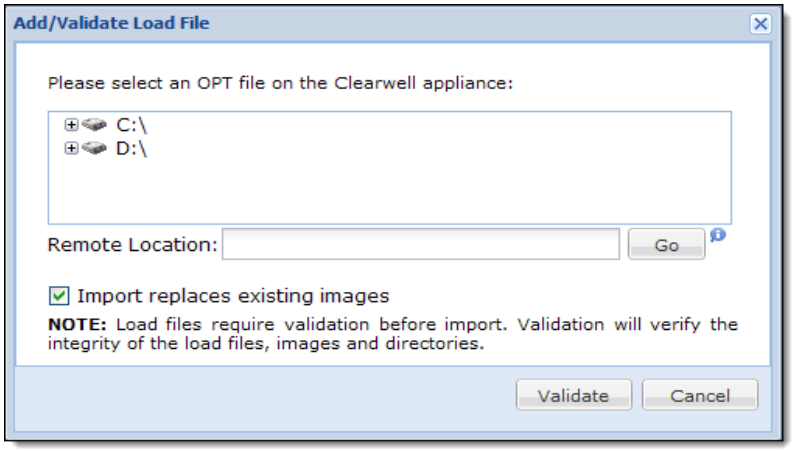
Column Number	Field Name	Description
– 1	Doc-id	Document ID of the item. For example: • For emails, it adheres to the format: 0.7.768.29275 • For attachments, it adheres to the format: 0.7.768.29275-000001
– 2	Volume Identifier	Not used. Can be empty (intentionally blank)
– 3	Image relative path	Path of the image relative to the location of the opticon file
– 4	Document break	**Y** indicates the start of a unique item. Should be blank otherwise
– 5	Folder break	Not used. Can be empty (intentionally blank)
– 6	Box break	Not used. Can be empty (intentionally blank)
– 7	Page count	Number of pages that the image represents

To import Native Images

Before proceeding with the import, review your load files to ensure they include the proper fields, syntax, and are in the Opticon format.

1. Go to the **Import Native Images** menu and click **Add/Validate Load File.**

   The **Add/Validate Load File** menu displays.



2. Provide the location information for the OPT load file. You can either select load files on the appliance or type the network location in the Remote Location text box and click **Go**.

   Note: You should have access to network locations and directories.

3. Click **Validate**. This verifies the integrity of the load file, images and directories and checks to see if the image load file will run successfully.

4. Once validation is complete, check the results under the Import Status column:

   – Ready for Import— Image load file passed validation. Proceed to Step 5.

   – Ready for Partial Import—Image load file completed validation but some items have errors. You have the following options:

   › Skip the items with errors and proceed to Step 5.

   › Fix the errors and then select revalidate icon to re-validate.

   – Failed—Items failed validation. Select the report icon to view the report for details of where the validation failed.

5. Select the import icon to begin the import. The Import Native Images displays the results of import actions.

   The Image Import Warnings menu displays confirming the number of items to be imported.

6. Click **Import Images** to proceed or **Cancel** to abort the import.

   Upon completion, check the import status column to confirm that is was successful. You will see one of the following states:

   Success—All items that passed the validation were imported and native image generation for each was successful.

   – Partially Finished—One or more items that passed the validation failed the native image generation.

   – Failed—None of the items that passed validation were imported.

7. View and monitor the results for the import action(s) on the **Import Native Images** menu.

Import Native Images

Field	Description
#Items	Total number of items in the Opticon file
Imported	Number of imported native items from the Opticon file
Replaced	Number of cached imported items that were replaced by a subsequent import
Errors	Number of import native item errors. See log file for error details.
Skipped	If the original item was in a production folder (locked) or has been redacted, they cannot be replaced. The native item import skips them.
Import Status	Native item import job status. When the validation job is running, the status will be "Validating" and when the import job is running, the status will be "Importing". In both the cases a value will be shown besides the status text to indicate a job in progress. Once validation completes the status can be: "Ready for import"," Ready for partial", and "Failed". During import the status can be: "Success", "Partially Finished" and" Failed

Generating Processing Reports

As **Case Admin**, you may need to generate reports to have an audit trail of the processing operation. Or you may want to show what files have been excluded from processing and track duplicate documents throughout an entire data set or on a per custodian basis. The central reports menu lets you generate these and other key reports.

Where Can I Find Processing Reports?

Starting with release 7.1.3, the reports formerly categorized under the **Pre-Processing Reports** tab are now located under the Processing module on the **Reports** tab (as of 7.1.3, the **Pre-Processing Reports** tab is deprecated). There are new reports and several of the pre-7.1.3 reports have been renamed or combined into a report that makes more sense.

Considerations

- For information on the available processing settings, refer to the "Configure processing parameters and features" section in the table "New Case: Processing Settings".
- Gain insight and verify file integrity by comparing the results for File ID, Strong File Type, and File Extension. You can do so in the following reports:
  - Not Processed Documents
  - Other Type - Extensions
  - Processing Reconciliation

  Note: For Container File ID information, see *"Supported Container Extraction File Types" on page 262*.

Some of the reports have been renamed. The following table identifies the new names for the pre-7.1.3 reports.

Processing Reports

Report Name	Description	Pre-7.1.3 Report Name
De-duplication	Generates a complete and comprehensive list of every location for each duplicated document. Filter by ALL, Case Folder, or Custodian. Filtering by Case Folder can be done at the source or sub-source level. **Note:** This report is not compatible with cases created prior to version 4.0.	Not applicable
De-duplication by Custodian	Generates a summary of duplicate document counts that are filtered on a per custodian or all custodian basis.	Not applicable
Discovery and Processing Options	Generates a list of the Discovery and Processing options **Note:** This report is not available for cases processed before 7.1.3.	Not applicable

Processing Reports (Continued)

Report Name	Description	Pre-7.1.3 Report Name
Discovery Errors	Generates a list of documents that resulted in Discovery errors **Note:** Run this report after processing for the most current list of file errors.	"Pre-Processing Errors" report
Not Processed Documents	Generates a list of documents that were processed but not discovered, or were excluded from Discovery and Processing (for example, items selected or excluded based on pre-processing options). **Notes:** • For loose message or EML files, the **Last Modified Time** column displays the sent date for the message. If no sent date is found for the message, this column is intentionally left blank. • Checking the NIST option produces the same NIST information as the pre-7.1.3 "Pre-Processing Excluded Known Files" report. • The **Container Count** column will be empty for pre-7.1.3 cases. • The checkbox option "**Include documents that are 'selected to process' but no yet processed**" option lets you toggle between processed data and data that is discovered but not yet processed.	"Pre-Processing Excluded Known Files" report (NIST list) "Pre-Processing Not Processed" report
Load File Discovery Errors	Generates a list of documents that resulted in Discovery errors during Load File Import For more information, refer to the *Load File Import Guide*. (Shows loose files only.)	"Pre-Processing Load File Import" report
Other Type - Extensions	Generates a summary and list of documents that are considered "Other Types" in Pre-Processing. This report can be generated at the source or sub-source level. When a source is selected, the report is generated for all its sub-sources. To generate the report at a sub-source level, only select the corresponding sub-source(s) and not its parent source.	Not applicable

Processing Reports (Continued)

Report Name	Description	Pre-7.1.3 Report Name
Processed Documents	Generates a list of the documents (PST, NSF, loose files) that were processed	"Already Processed Files" report

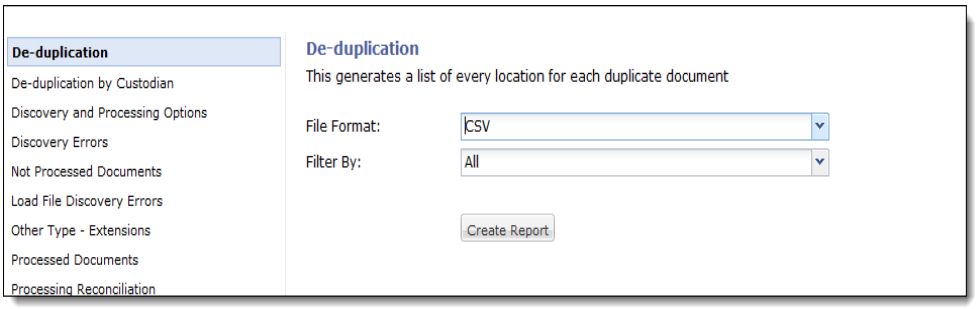
Processing Reports (Continued)

Report Name	Description	Pre-7.1.3 Report Name
Processing Reconciliation	Generates a summary and list of files on disk and the corresponding document and items counts needed for reconciliation. This report can be generated at the source or sub-source level. When a source is selected, the report will be generated for all its sub-sources. To generate the report at a sub-source level, only select the corresponding sub-source(s) and not its parent source. Consult the ReadMe.txt file included in the report zip file for more information about the layout and meaning of the various fields. **Note:** This report is not available for cases processed before 7.1.3. **Considerations**: • Accurate reporting: Keep in mind that if data is being processed while running the Processing Reconciliation report or if data is processed after the report has been run, the report will not reflect the latest state. If data has been processed after or while the report was being run, rerun the report to ensure the latest information is included in your report. • Any invalid characters in the directory structure that is created for the report are replaced with equivalent Unicode characters. This replacement also avoids collisions between two directory names. The escape character used is '@'. • Any file names that have invalid Windows characters are replaced with '_' (underscore). • The Associated File column displays whether the file is associated with another file or not. Examples of associated files include TIFF/text pairs, EML/EMLXPART files, and LEF/L01 files. For TIFF/text pairs, text is accounted for under the Associated File column and its TIFF counterpart is listed under Documents Processed.	Not applicable

To generate reports

1. Under the **Processing** module for a selected case, click **Reports**.

   The Reports screen appears and lists available reports.



2. Choose one of the following report options:

   Note: For additional report information, see *"Processing Reports" on page 99* and *"Step 7: Review Processing Results" on page 157*.

   - **De-duplication**
   - **De-duplication by Custodian**
   - **Discovery and Processing Options**
   - **Discovery Errors**
   - **Not Processed Documents**
   - **Load File Discovery Errors**
   - **Other Type - Extensions**
   - **Processed Documents**
   - **Processing Reconciliation**
   - **Processed Audio Size and Duration** (requires Audio Search Module)

3. Specify the information in the following table that is relevant to the report.

Report Settings

Field	Description	Report or Comment
File Format	Specify the output format for the report (CSV or XLSX). **CSV option is available for:** • Deduplication • Discovery Errors • Not Processed Documents • Load File Discovery Errors • Not Other Type - Extensions • Processed Documents **XLSX option is available for:** • Deduplication by Custodian • Discovery and Processing Option	XLSX produces a very nicely formatted file but may take longer to generate, especially for reports that generate large lists of files from a large dataset. The CSV option is faster but simpler.
Filter By	Specify the filter for the report data: (All, Custodian, or Case Folder). "All" is the default.	Control what information is included in the report by designating a filter from the list.
Batch Label	Select a label type from the drop-down list, or "All Labels" is applied by default. The reports that use this setting are: • De-duplication by Custodian • Discovery Errors • Not Processed Documents • Processed Documents	These reports are useful in analyzing and preparing your data for processing, especially when the data can be easily identified by batch. Select the batch label and run a report for that specific batch. **Note:** Batch-level reports are available for processing batches in cases created prior to version 6.6. For cases created in 6.6 and later versions, additional reports for discovery batches are also available.
Select Type:	Select one of the options ("By Discovery", "By Processing" or "By Sources")	For "Discovery and Processing Options" report
Discovery Time	Select a label from the drop-down list, or "All Labels" is applied by default.	For "Discovery and Processing Options" report
Reason Code	Select a reason that the document was not processed from the drop-down list. The default is "All Errors".	For "Discovery Errors" Report
Exclusion Criteria	Select one or more exclusion criteria from the list.	For "Not Processed Documents" Report

4. Click **Create Report**.

The report is generated and the job becomes available in the Jobs window for download.

Note: If you need to submit multiple jobs of the same report, you can do so without waiting for the previous job to complete.

Processing Source Data

After you have verified that your sources have been added correctly, and have been pre-processed, you can process your source data.

Note: Depending on your processing requirements, you can use various support features to improve your processing efficiency. If your source files include encrypted data, provide the passwords or certificates to decrypt the data during processing. If you anticipate processing large amounts of cached data, administrators can configure the storing of cached data in another location, either on or off the appliance. Use the Support Features screen to apply appropriate properties for your processing needs. For more information, refer to *"Using the Support Features" in the System Administration Guide*.

To process your source data

1. On the top navigation bar, for a selected case, click **Processing > Sources and Pre-Processing**.
2. From the **For Selected Items** menu, select either **Start Processing Source without Discovery** or **Start Processing Source with Discovery**.
   - If no additional files have been added to the source since it was added, select Start Processing Source without Discovery.
   - If additional files have been added, Start Processing Source with Discovery must be used to discover the newly added files.
3. Click the **Go** button to start the selected task.

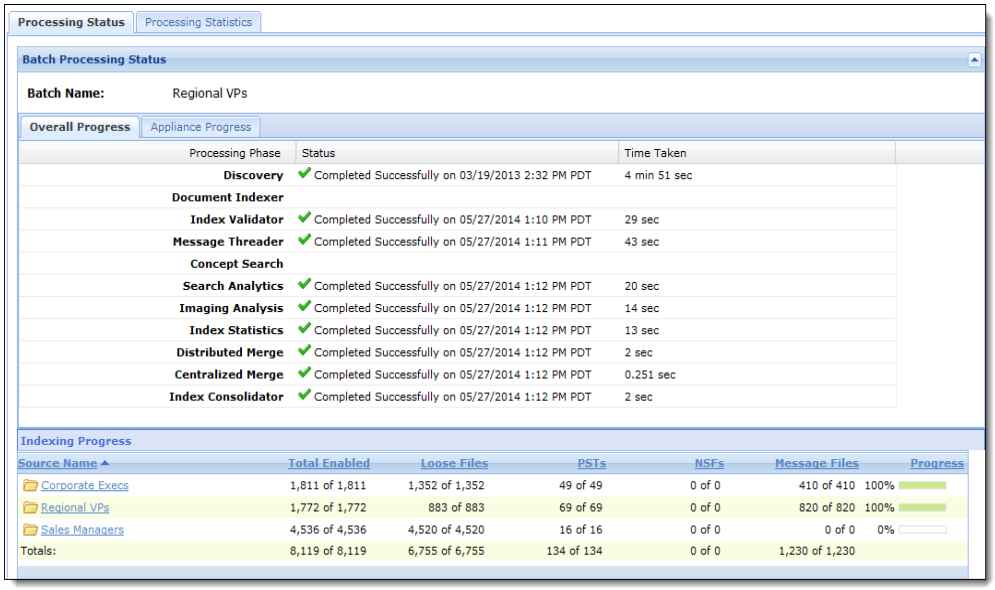
Monitoring Source Processing Status

The Processing Status screen displays the status of the Indexer and Analytics services, statistics on the collected and indexed content, and the indexing progress for each document source in the case.

To view the processing status

1. On the top navigation bar, select a case, then click the **Processing** module.

   The Processing Status screen displays.



2. Under the **Processing Status** tab, the Batch Processing Status section displays the name of the processed batch and percent complete. Two status views are available:

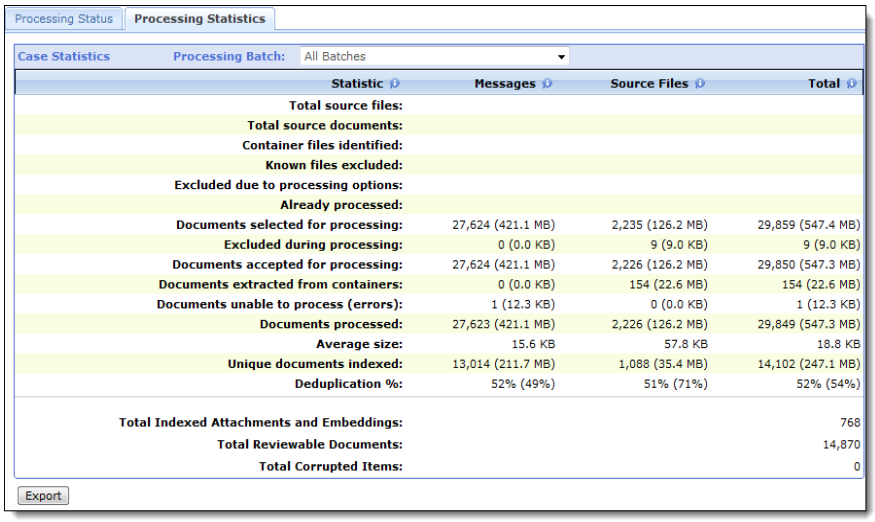
   A. The **Overall Progress** sub-tab provides a high-level view of each processing phase and current status of the case. (If any problems occurred during the last processing cycle, click the warning ⚠ icon next to the source for more details.)

   The Status column indicates whether each service is running, has successfully completed, or has never been started. The Time Taken column shows how long each phase of processing took to complete.

   B. The **Appliance Progress** sub-tab shows which appliance completed the processing job, and the current status of the appliance. (If you have more than one appliance provisioned for distributed processing, multiple appliances will be displayed.

   To view indexing detail, scroll down to the **Indexing Progress** section. This area lists each document source, the total number of items enabled for indexing, the number of individual files, such as Loose Files, PST, and NSF files, the number of message files, and the percentage of the collected content that has been indexed. To view or change a source's configuration, click the source name (see *"Managing Case Sources and Custodians" on page 49*).

3. Click the **Processing Statistics** tab to view statistics of the case after processing.



Statistic	Messages	Source Files	Total
Total source files:			
Total source documents:			
Container files identified:			
Known files excluded:			
Excluded due to processing options:			
Already processed:			
Documents selected for processing:	27,624 (421.1 MB)	2,235 (126.2 MB)	29,859 (547.4 MB)
Excluded during processing:	0 (0.0 KB)	9 (9.0 KB)	9 (9.0 KB)
Documents accepted for processing:	27,624 (421.1 MB)	2,226 (126.2 MB)	29,850 (547.3 MB)
Documents extracted from containers:	0 (0.0 KB)	154 (22.6 MB)	154 (22.6 MB)
Documents unable to process (errors):	1 (12.3 KB)	0 (0.0 KB)	1 (12.3 KB)
Documents processed:	27,623 (421.1 MB)	2,226 (126.2 MB)	29,849 (547.3 MB)
Average size:	15.6 KB	57.8 KB	18.8 KB
Unique documents indexed:	13,014 (211.7 MB)	1,088 (35.4 MB)	14,102 (247.1 MB)
Deduplication %:	52% (49%)	51% (71%)	52% (54%)
Total Indexed Attachments and Embeddings:			768
Total Reviewable Documents:			14,870
Total Corrupted Items:			0

Note: Move the mouse over a cell in any column to display more information about its content.

This view includes the following processing detail:

A. The **Statistic** column provides case statistics for the document crawler and indexing services. Choose a specific batch job or **All Batches** to view the results for all jobs. These statistics are cumulative, and are not reset when the appliance is rebooted.

B. The **Messages** column includes email documents that may be processed from an email server or archive or from an individual email file (.msg/.eml). It also includes non-email items such as contacts and calendar entries.

C. The **Files** column includes counts for any non-email documents, excluding email attachments.

D. The **Total** column combines the total numbers for each row in the previous columns for each statistic.

To view the remaining disk space on an appliance, refer to the section *"Maintaining eDiscovery Appliances" in the -System Administration Guide*. The following table describes each of the statistics. Hover over any entries in the list to view additional information.

Note: To view file and message warnings, and unprocessed documents or mailboxes, see the **Processing > Exceptions** screen.

Refer to the following table for full detail of each statistic type:

Processing Statistics

Statistic	Message	Files	Total
Total source files		Total source files (including loose files and email containers such as PSTs and NSFs) when this batch was processed	Total source files (including loose files and email containers such as PSTs and NSFs) when this batch was processed
Total source documents	Total messages after extraction from email containers (PSTs/ NSFs)	Total files after separating out emails and email containers as messages	Total documents found in source locations, including loose files and individual emails
Container files identified		Loose files identified as containers	Loose files identified as containers
Known files excluded		Known files excluded due to known file filtering (also known as "de-NIST-ing")	Known files excluded due to known file filtering (also known as "de-NIST-ing")
Excluded due to processing options	Messages excluded from processing based on user-selected processing options	Loose files excluded from processing based on user-selected processing options	Total documents excluded due to user-selected processing options
Already processed	Messages that have already been processed in previous batches	Loose files that have already been processed in previous processing batches	Total documents already processed in previous processing batches
Documents selected for processing	Messages selected for processing	Loose files selected for processing	Total documents selected for processing
Note: The count and volume of documents selected for processing are calculated after extraction, and will almost always be higher or lower than the count and volume of documents in the original source location (sometimes substantially), depending on the compression ratios of the PST, NSF, and container files.			
Excluded during processing	Messages and non-email message documents (like contacts) excluded during processing	Loose files excluded during processing	Total documents excluded during processing
Documents accepted for processing	Messages accepted for processing (selected minus excluded)	Loose files accepted for processing (selected minus excluded)	Total documents accepted for processing (selected minus excluded)

Processing Statistics (Continued)

Statistic	Message	Files	Total
Files extracted from containers	Messages processed that were extracted from archive containers (such as ZIP files)	Loose files processed that were extracted from archive containers (such as ZIP files)	Total documents processed that were extracted from archive containers (such as ZIP files)
Documents unable to process (errors)	Messages that could not be processed due to an error condition (for example, the crawler dropped due to lack of PKI certificates)	Loose files that could not be processed due to an error condition (container files with error during discovery) Unprocessed documents due to case exceptions	Total documents that could not be processed due to an error condition
Documents processed (Remove "(post-extraction)" text)	Messages processed by the product (prior to deduplication)	Loose files processed by product (prior to deduplication)	Total documents processed by product (prior to deduplication)
Average size	Average message size (including attachments) across all messages processed	Average loose file size across all loose files processed	Average document size across all documents processed
Unique documents indexed	Unique messages indexed (after deduplication)	Unique loose files indexed (after deduplication)	Total indexed documents (this will match the number of documents returned in an "empty search")
Deduplication %	Percentage by which processed message count (volume) was reduced by deduplication	Percentage by which processed loose file count (volume) was reduced by deduplication	Percentage by which total processed document count (volume) was reduced by deduplication"

4. To export the information on the Processing Status screen, click **Export**.

Viewing Processing Exceptions

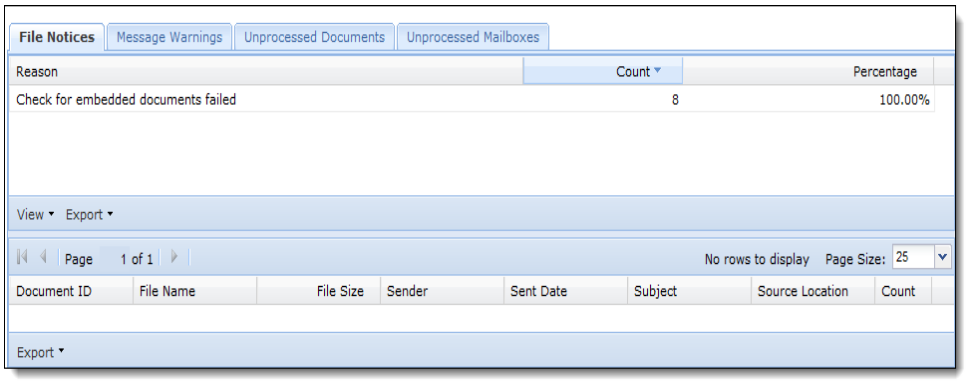
The system reports on the following types of exceptions that may occur during processing:

- **File Notices**—File issues discovered during processing, such as empty or corrupt files. Includes the reason and the number and percentage of files.
- **Message Warnings**—Exception messages about the documents that were processed, such as a missing sent time in email messages, or errors in processing attachments. Includes the reason and the number and percentage of files.
- **Unprocessed Documents**—Documents that could not be processed. Includes the reason and the number and percentage of files.
- **Unprocessed Mailboxes**—Mail boxes that could not be processed. Includes the mail box name, mail server, reason, label, and start and end times.

Note: Information messages (for the first three exception types) are hidden by default. To view warning-only exceptions for case users, see step 2 in the following procedure. You can change the default setting for this option by using the Property Browser in Support Features. For details refer to the section *"Using the Support Features" in the System Administration Guide*.

To view and save information on the processing exceptions

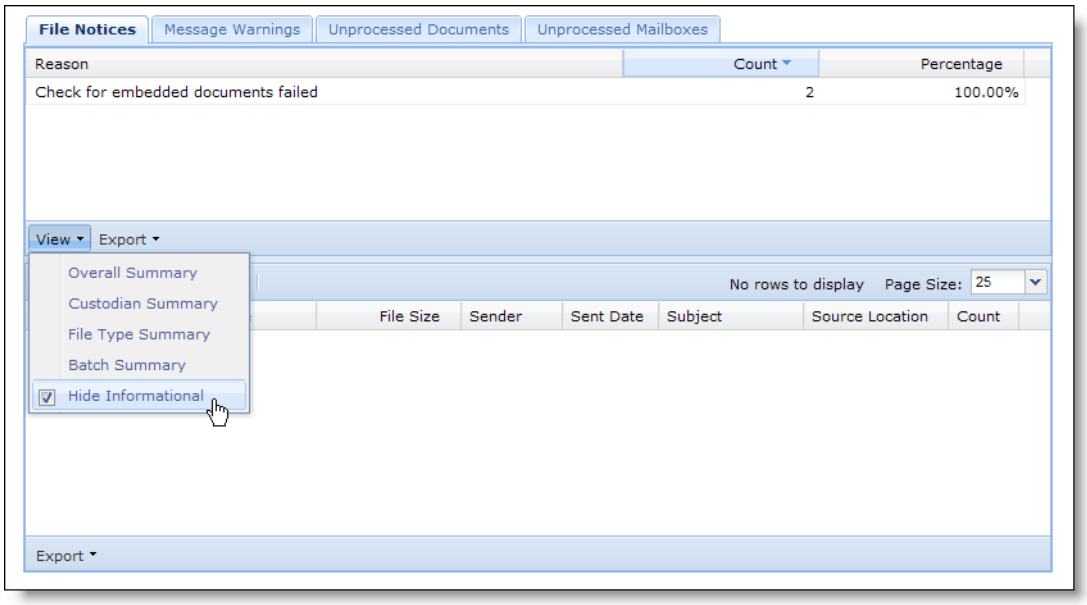
1. On the navigation bar, for a selected case, click **Processing > Exceptions**.



Note: For a full list of all possible exceptions messages for file notices and message warnings (both errors, and informational), see *"Processing Exceptions" on page 159*.

2. Select the tab for the type of exceptions you want to view.

   Note: To view all exceptions, or detailed information (not just warnings), click the **View** drop-down menu and clear the **Hide Informational** check box (selected by default).



3. Each tab indicates at a high level the types of issues that were encountered.

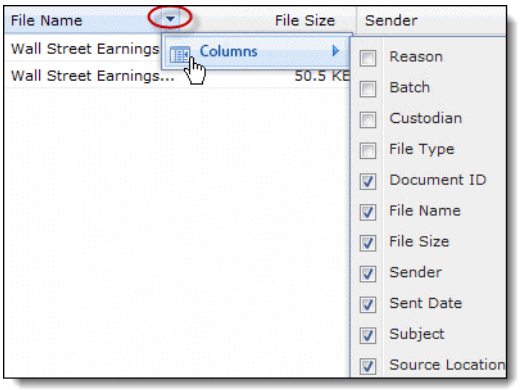
   Note: You may notice a difference in exception counts between the Exceptions screen and filters in your search results. This is because the Exceptions screen displays *every occurrence* of a file, while search results and filters account for *unique* files *only once*. For example, if an exception appears in a loose file, and that same loose file is in the source data five times, the count on the Exception screen will display "5" (for every occurrence of that file). However, in your search results, that same file is only indexed once (with deduplication), and thus will display "1" (for that file).

   To view details for an issue type, highlight that entry in the exception list, click **View**, and choose from the following options:

   - **Overall Summary**—General summary of the information.
   - **Custodian Summary**—Information organized by custodian.
   - **File Type Summary**—Information organized by file types.
   - **Batch Summary**—Information organized by batches.

4. To export information on a selected type of exception, click **Export** and choose from the following options:
   - **Export Current List to CSV—**Export the currently-displayed data to a CSV file.
   - **Export All to CSV—**Export all of the exception information for the case to a CSV file.

   Export current and all document options can select any of the locations present in the export location menu. The list of available locations depends on whether group access membership and location is configured or if external export locations have been configured with the property browser.
   - **Export Current Documents—**Export the documents that are involved in the currently-display exceptions. Select the location.
   - **Export All Documents—**Export all of the documents that are involved in exceptions for the case. Select the location.
5. To add or remove columns from the tables on this screen, click the down facing arrow for the column, choose **Columns**, and select the check boxes for the desired columns.



Processing (or Resubmitting) Documents for OCR

At case setup, it is recommended that you do not enable OCR processing, so as to maximize processing time and prevent a slowdown of system performance. Now that your case data has initially processed, you have the option of running those documents through OCR for the first time, as a way to "resubmit" those documents into your processed case data.

You may also have identified additional documents you want processed with OCR. In either case, documents can only be processed one time, and cannot be re-processed. Note that processing with OCR will take considerably longer than regular processing time.

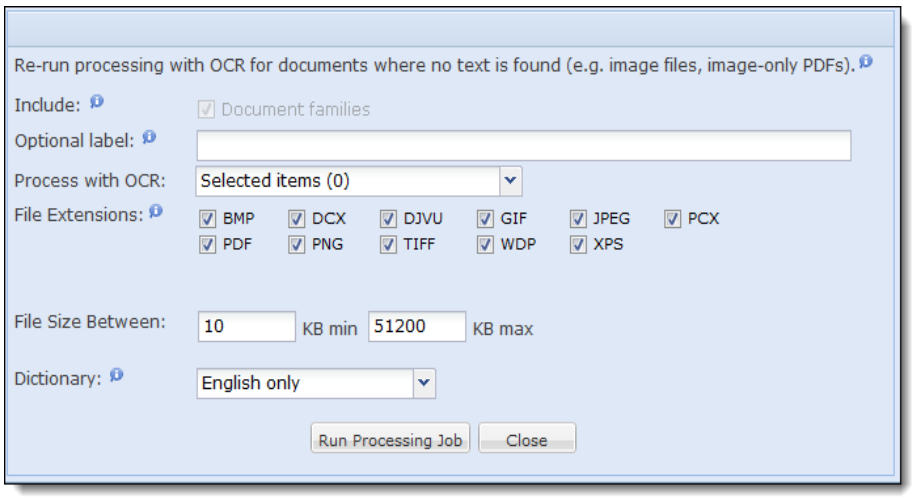
Note: Users must have the **Case Manager**, **Case Admin**, **Group Admin**, or **System Manager** role to resubmit documents for OCR processing. This allows you to perform OCR processing on a set of search results. Documents to be OCR-processed (typically in a folder search) are selected from a set of search results. Follow the steps in this section to resubmit a batch for OCR.

CAUTION: Ensure your case is not currently in active review. This could cause a serious impact for all other users currently accessing or performing reviewer actions on this case.

To process documents for OCR (after initial case processing)

1. From the **Analysis and Review** module, with your search result set open on the Documents screen in List view, click **Actions > OCR**.

   The OCR window appears:



2. Check the document families check box to include all document families across the case that are related to the selected items. In this example, all the document families have already been brought in.
   - A document family is the parent document and all attachments.
   - An item is any individual email message, attachment, embedding, or other individual piece of content.

3. (Optional) Type a description/label for this OCR batch. The start and end time are automatically associated with the batch.
4. Select which documents you want to process with OCR. If you pre-selected a batch, the **Selected messages** option showing the total selected appears by default. Alternatively, select **Entire search results** (showing the total documents in your result set) or **Selected items** (showing the total items).
5. Choose one or more file extensions to be processed with OCR. By default, the file types selected reflect those in your case settings ("OCR Processing") configuration.

   Note: The product identifies files to be processed with OCR based on the selected file extensions. If you select "PDF", PDF files will be recognized regardless of file extension.
6. Select a range for the file size (minimum KB to maximum KB).
7. Select the dictionary for the language you want to use. The OCR engine will try to recognize characters from all selected dictionaries.

   Note: You cannot reprocess the same documents with OCR more than once. However, if any documents failed to process the first time, those can be tried again until they process. If you run OCR processing with one language dictionary, and later discover other languages in the batch, you will not be able to reprocess those with the appropriate dictionary. However, the product attempts to recognize characters from all selected dictionaries.
8. Click **Run Processing Job**.

To view processed jobs

1. Once your OCR job has processed, view the results.
   - Click the Jobs window above the navigation bar.
     The Jobs window shows the job currently running, before it completes. To stop the job before it completes, click the cancel icon in the Actions column.
   - To view the status and details, from the **System** view, click **Jobs**.
2. From either view, click the icon under the Status column to view details in the Job Status Log.

To search for OCR processed files

Before you begin: After your OCR job has processed, you can search on the results.

1. From the **Analysis and Review** module, click **Advanced Search**.
2. From the Advanced Search screen, click the Identifiers section and select an option to **Match** [**Any**] **of these OCR batches** (that have recently been processed).

   For more information about this Advanced Search (OCR) option, refer to the table in the section *"Standard Advanced Searches" in the Veritas eDiscovery Platform User's Guide*.

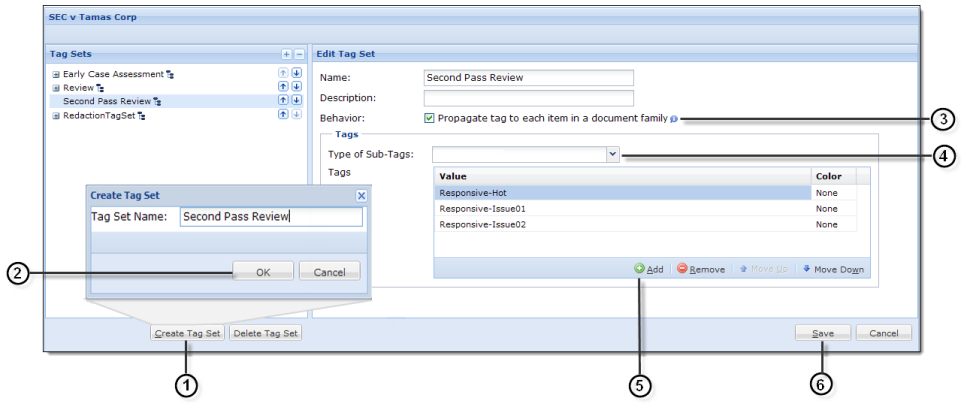
Defining Tag Sets

Tag sets are identifiers with yes/no (check box) or predefined values that users can select to indicate the status (or other information) of the documents in a case. For example, you can define an "Assigned To" tag with a list of reviewer names, and a "Reviewed" check box tag that can be set after a document is reviewed. You can define up to 100 tag sets.

You must have the **Case Manager**, **Case Admin**, **Group Admin**, or **System Manager** role to define tag sets and tags.

To add a tag set

1. With a case selected, click **Analysis and Review > Tags**, then click **Create Tag Set**.



2. On the Create Tag Set dialog, type the Tag Set Name and click **OK**.
   The tag set name displays in the Tag Sets pane.

3. In the Edit Tag Set pane, select whether you want the tag to be propagated to each item in a document family.

   Note: If enabled, all item tags are propagated to each item in the document family. For example, an email and its attachment receive the same tag when either is tagged. If not enabled, each item must be tagged separately.

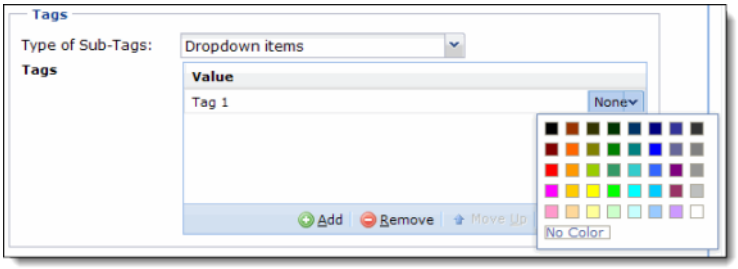
4. Select the type of Sub-Tags from the drop-down menu.

5. On the Tags table, click the **Add** button.

6. Provide a name for the tag and click **Save**.

Repeat this process for each new tag you want to add.

Tag Settings

Field	Description
Name	Enter a tag category name (up to 31 characters).
Description	Enter a tag category description (up to 255 characters).
Tag Category Type	Select whether the tag set uses check boxes, radio buttons, or drop-down list with selectable values. When you select an option, the area expands to allow you to define specific values.
Tags	Click **Add** to add a value (up to 32 characters). Enter the value name. To add a color to the value, click **None** on the right side of the entry, and select a color.  To move a tag value up or down one position on the list, select the value and click the up or down arrow keys. Tag values are displayed to the user in the sequence shown here. To delete a tag value, select the value and click **Remove**. Note that any saved searches that include the tag value will also be deleted.

To create sub tags for an existing tag value

1. On the top navigation bar, for a selected case, click **Analysis and Review > Tags**, then select the tag that you want to add a subtag to.
2. Select the tag value in the tag list to display the nested tag settings in the **Edit Tag** area on the right.
3. Select the type of subtag (check box, radio button, or drop-down list).
4. Select whether subtagging is optional, recommended when the tag is selected, or required when the tag is selected.
5. Add, move or remove tag values as you would for any other tag set.
6. If you want to prepend text to the subtag, select **Prepend to sub-tag value** and select either the tag name or an optional value.
7. Select whether reviewer comments are optional, recommended when the tag is selected, or required when the tag is selected, and enter a label for the reviewing comments entry.
8. Click **Save** to save the values for the subtag, or click **Cancel** to discard your changes.

To edit an existing tag set

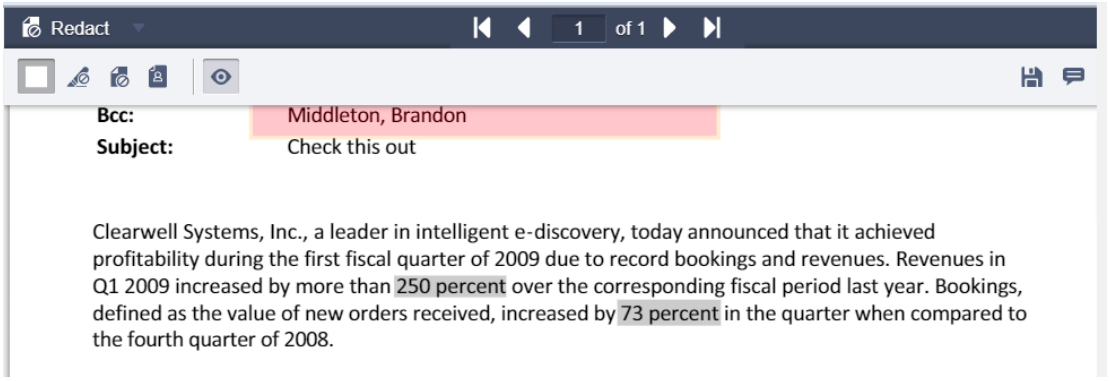
1. On the top navigation bar, for a selected case, click **Analysis and Review > Tags**, then select the tag that you want to edit.
2. To change a tag set name, description, or list of tag values, click the tag set name, enter your changes, and click **Save**.
3. To move a tag sets up or down one position on the list, use the arrows on the right of the **Tag Sets** list. Tag sets are displayed to the user in the sequence shown here.
4. To delete a tag set, select the set, and click **Delete Tag Set**. Click **OK** to confirm. The deleted tag category is removed from all documents, and any saved searches that include the tag category are also deleted.
5. Click **Save** to save the changes.

Additional Configurations for Redactions

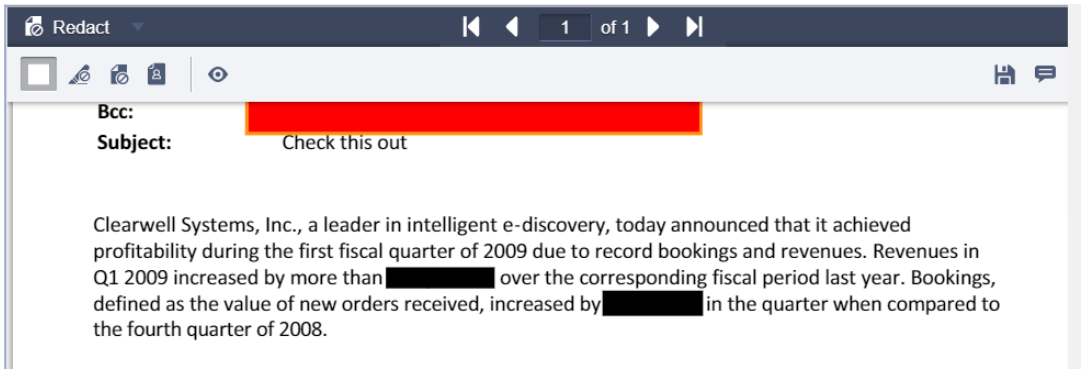
Configuring default redaction view mode

Administrators can control the default redaction style, opaque or translucent, for redaction view mode. The default redaction view mode is translucent.

Default translucent redaction view mode:



Opaque redaction view mode:



The following property can be set using **System** > **Support Features** > **Property Browser** to configure the default redaction style:

Property: ***esa.prizmdoc.redaction.view.default.transparency.mode***

Values: **Normal** - For Opaque redactions and **Draft** - For translucent redactions

Configuring number of retries for bulk redaction jobs

By default, the number of retries for a bulk redaction job is configured as 3. An administrator can configure the number of retries by configuring the following property using Property Browser.

Property: ***esa.bulkredaction.prizmdoc.retry.attempts***

Value: Any integer

By default, number of retry attempts is set to 3.

Configuring length of preset reason codes

An administrator can configure the maximum number of characters allowed in each preset reason code of a redaction set using the following property.

Property: ***esa.case.reason.code.length.maxlength***

Value: Any Integer

Default value: 40

Configuring default review mode

Starting with release 10.0, the default view in review mode on the Document Review Screen is Native/Image view. An administrator can configure the following property to set the default review mode.

Property Name: ***esa.default.view.in.review.mode***

Values:

native - For Native view

html - For Text view

Default value: **native**

Configuring font family for header, footer, and watermark

By default, the Times New Roman font is used to specify header/footer and watermark text in Analysis & Review > Folders > Production Folder > Header/Footer tab.

An administrator can use the following properties to specify the font for header, footer, and watermark for a specific case (case-level) or all cases (system-level) using **System** > **Support Features** > **Property Browser**.

Property Name:
esa.production.folder.header.footer.textfield.length.autoCalculateUsingFontSize

Description: Determines whether the the header/footer text field length should be auto calculated using the font size specified for header/footer in a Production Folder.

System/Case-level Property: System-level

Default Value: TRUE

Property Name: ***esa.production.folder.header.footer.textfield.length***

Description: Specifies the header footer text field length when the length is not calculated using the font size.

System/Case-level Property: System-level

Default Value: 30

Property Name: ***esa.production.folder.watermark.fontFamily***

Description: Specifies the default font for the Production Folder watermark text for a specific case.

System/Case-level Property: Case-level

Default Value: Times New Roman

Property Name: ***esa.production.folder.watermark.fontFamily.global***

Description: Specifies the font for the Production Folder watermark text for all cases.

System/Case-level Property: System-level

Default Value: Times New Roman

Property Name: ***esa.production.folder.watermark.fontWeight***

Description: Specifies the font weight for the Production Folder watermark text font
Valid values are normal and bold.

System/Case-level Property: System-level

Default Value: normal

Property Name: ***esa.production.folder.watermark.opacity.percentage***

Description: Specifies the Production Folder watermark text opacity percentage rage. Valid values range 1 to 75. Greater the value, lesser is the transparency.

System/Case-level Property: System-level

Default Value: 20

Property Name: ***esa.production.folder.watermark.fontSize***

Description: Specifies the font size for Production Folder watermark text.

System/Case-level Property: System-level

Default Value: 150

Property Name: ***esa.production.folder.watermark.rgb.textColor***

Description: Specifies the Production Folder watermark text color in r,g,b format. The value of each r/g/b component should be between 0 to 255.

System/Case-level Property: System-level

Default Value: 0,0,128

Property Name: ***esa.production.folder.header.footer.rgb.textColor***

Description: Specifies the Production folder Header/Footer text color in r,g,b format. The value of each r/g/b component should be between 0 to 255.

System/Case-level Property: System-level

Default Value: 128,128,128

Property Name: ***esa.production.folder.header.footer.fontFamily***

Description: Specifies the font for the Production folder header text for a specific case.

System/Case-level Property: Case-level

Default Value: Times New Roman

Property Name: ***esa.production.folder.header.footer.fontFamily.global***

Description: Specifies the font for the Production folder header text for all cases.

System/Case-level Property: System-level

Default Value: Times New Roman

Property Name: ***esa.production.export.transparent.redaction.opacity***

Description: Specifies the Production export transparent redaction opacity percentage. The valid values range 1 to 75. Greater the value, lesser is the transparency.

System/Case-level Property: System-level

Default Value: 20

Property Name: **esa.export.transparentRedactions.enabled**

Description: Controls the option to burn transparent redactions during export.

The valid values are:

true - Production Export UI will show a check box to enable/disable burning transparent redactions during export.

false - Production Export UI will hide the check box to enable/disable burning transparent redactions during export.

System/Case-level Property: System-level

Default Value: true

Setting Up Folders

Folders are an administrative aid used to organize sets of documents or items. The folders are presented as options when performing searches, and are subject to searching and configuration rules. You can create non-production, production, and review set folders.

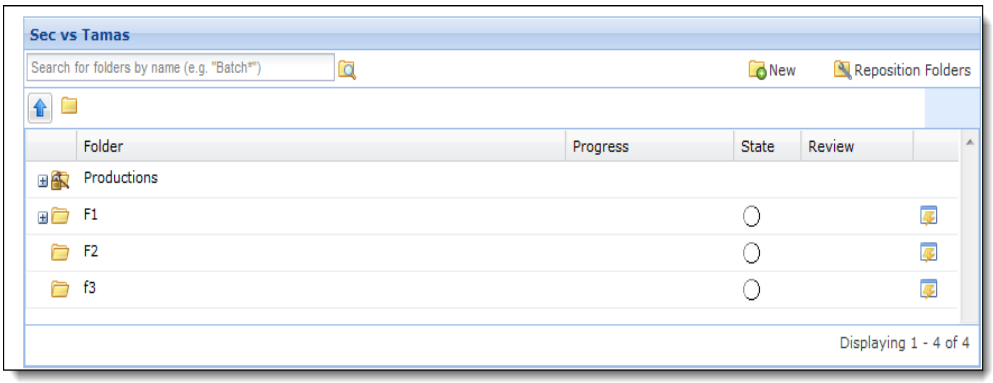
The granularity of item-level searches and views allows you to easily select only items of interest (or exclude the entire document family) for review folders or for a custom folder workflow. Of course, you can always select all "hit" items and the parent document families to copy to a folder.

You must have the **Case Admin**, **Group Admin**, or **System Manager** role to set up folders.

Set Up Non-Production Folders

To set up standard (non-production) folders

1. On the top navigation bar, for a selected case, click **Analysis and Review > Folders**.



2. The list shows each folder. To expand a folder and show the associated subfolders, click the + sign to the left of the entry. Click the - sign to collapse the entry.
3. To edit an existing folder:
   A. Select the folder and click **Edit Folder** icon.
   B. Modify the folder name or optional description, and click **OK**.
4. To add a new folder;
   A. Select the folder just above or below the one you want to add, or select the folder for which you want to create a sub-folder.
   B. Click **Add Folder** and choose the **Above**, **Below**, or **As Subfolder** option.
   C. Enter a folder name or optional description, and click **OK** to add the new folder.
5. To rearrange folders, select the **Reposition Folders** check box and then drag and drop the folders to their new locations in the list.
6. To remove a folder, select it and click **Delete Folders**. Click **OK** to confirm.

Creating Review Set Folders in Batches

You can divide your search results for easy management and distribution among the users reviewing the documents. Documents (whether at the item or document family level) are automatically grouped into folders, called review sets or batches, based on the criteria you specify.

You can create folders for review and specify that your search results are divided into a specific number of batches or that each review set, or folder, contain a certain number of documents. In addition to batching documents found by your search query, you can include discussion threads that expand the scope of your search.

Overview

Follow these steps to create review sets in batches.

1. (Optional) Set up your folder hierarchy manually.

   See *"Set Up Non-Production Folders" on page 124*.

2. Create multiple review sets.

   See *"Create Multiple Review Sets" on page 126*.

3. If necessary, set visibility to the review sets.

   See *"Set Visibility to Review Sets" on page 127*.

4. Assign access to the review sets to enable folder management tracking.

   See *"Assign Access to a Review Set" on page 127*.

See also folder management options

5. To change the name or description of a review set.

   See *"Edit a Review Set" on page 137*.

6. To keep organized, reposition or delete review folders.

   See *"Reposition or Delete a Review Set" on page 128*.

For more information on reviewer folder operations, refer to the section *"Review Set Management" in the User Guide*.

Create Multiple Review Sets

Before you begin: If possible, log in to an account that shares the same access profile as the Reviewers who need access to the documents.

Note: To create multiple review sets with the Batch interface, you must have the "Allow folder setup" permission set. The **Case Manager**, **Case Admin**, **Group Admin**, and **System Manager** roles have this permission.

To create multiple review sets from a batch

1. From the **Analysis and Review** module, run your search query.
2. From the **Actions** menu, click **Batch**.

   The Batch interface displays.
3. Type a new folder name or select the name of an existing folder where you want the review sets created. Search results (including their document families) are divided and split into folders called Batches.

   For example, if you want the review sets to be organized by your review process, a folder might be named "Early Case Assessment" or "Quality Control". These organizing folders must be created before you create your batches.
4. Type the batch folder prefix.

   Each review set uses this prefix.
5. Select how you want the documents divided.
   - **Number of batches.** Documents are equally distributed among the number of review sets, or batches, you specify.
   - **Documents per batch.** The number of review sets, or batches, is determined by the number of documents in your search results.
6. (Optional) Specify whether you want all discussion threads to be placed in the same review set, or batch. This may mean that the number of documents is not the same in each batch.

   Note: This option can be helpful for identifying themes within a document set.
7. (Optional) Specify whether all related discussion thread documents are included in the review set, or batch. This may mean that the total number of documents batched is greater than the number of documents in the search result.

   Note: This option increases the number of documents to be reviewed. Documents outside the parameters of your initial search query will likely add to the search results.
8. If necessary, select how additional discussion threads are handled.
9. Click **Create Batches**.

Set Visibility to Review Sets

Setting review set visibility enables reviewers to see specific review set folders and their contents. If the reviewers' access profiles are set to Show All Documents or you used the same Access Profile as your reviewers, reviewers will automatically have access to all review sets.

Before you begin: You only need to perform this procedure if you create a batch of review sets using a user account with an access profile that the Reviewers do not share AND the Reviewers have restricted access to folders.

Note: You must have the "Allow Case Management" and "Allow User Management" permissions to create new Access Profiles. The **Case Manager**, **Case Admin**, **Group Admin**, and **System Manager** roles have this permission.

To set visibility to review sets

1. On the top navigation bar, select the appropriate case from the drop-down menu and click **Case Home**.
2. Click **Users**.
3. On the Users screen, click the **Access Profiles** tab.
4. Select the access profile of the reviewers who need access to the review sets.
5. On the Edit Access Profile screen, click the **Documents** tab.

   The Restrict Visibility option should be selected.

   Note: If Show All Documents is selected, the review sets are currently visible to all reviewers with this access profile.
6. Select the **Show Folder and Contents** option for the specific review sets to be accessed.
7. Click **Submit**.

Assign Access to a Review Set

Assigning access to a review set enables a reviewer to use the review set management feature, a simple way to track who is reviewing which review set and what the current review set status is. After access is assigned, a reviewer can use the Manage Folders window to begin, stop, or complete the review of a review set.

Note: You must have the "Allow folder check-out management" permission to enable users to check in and check out review set folders. The **Case Manager**, **Case Admin**, **Group Admin**, and **System Manager** roles have this permission.

To assign access to a review set

1. From the **Analysis and Review** module, click the **Folders** drop-down menu.

   This menu is next to the search field and displays "All Documents" by default.

2. Click next to the folder you want to assign, then from the **Actions** menu select **Assign**.

3. On the Assign Review Access window, click and drag user names from the left "Available Case Users" box to the right to designate them as "Assigned Case Users".

4. When finished, click **OK**.

Reposition or Delete a Review Set

Repositioning, copying, or deleting a review set is the same as performing these functions on any folder. (Repositioning folders maintain their state after moving.)

Note: You must have the "Allow folder setup" permission to manage review set folders. The **Case Manager**, **Case Admin**, **Group Admin**, and **System Manager** roles have this permission.

To reposition or delete a review set

1. From the **Analysis and Review** module, click the **Folders** drop-down menu.

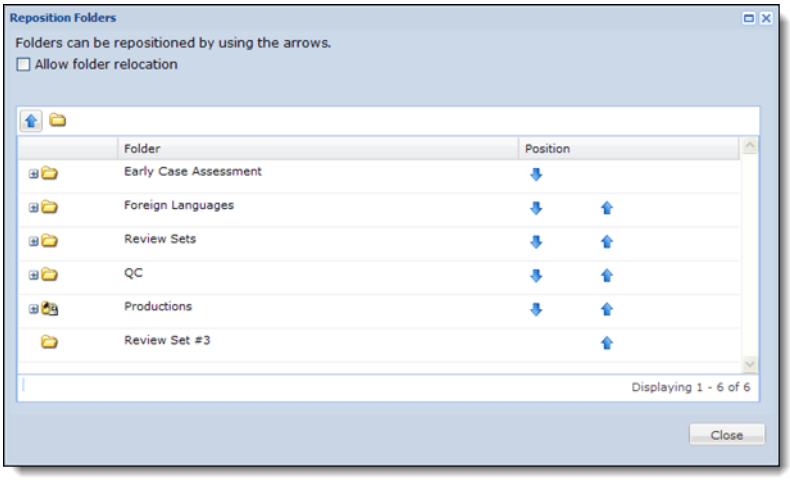
   This menu is next to the search field and displays "All Documents" by default.

2. Click to navigate to, or select the review set that you want to reposition or delete.

3. If repositioning a review folder:

   A. Click **Reposition Folders**.

   The Reposition Folders window appears.



Similar to the main folder management window, you can navigate to any subfolder and breadcrumbs will show your current location. (Click the arrow icon anytime you want to return to the home screen in this window.)

B. Reposition folders by using clicking the arrows to move them up or down. Click + next to any folder to reposition the its subfolders.

C. When finished, click **Close**.

4. If deleting a review folder:

A. Click the edit/review icon next to the folder you want to delete, then from the Actions menu select **Delete**.

B. On the confirmation dialog, click **OK** to continue deleting the folder. (Folder contents will not be deleted, but will remain in the "All Documents" collection.)

Setting Up Production Folders

The Production folder is a system-generated, top-level folder, under which you will create subfolders for your individual production. You can specify production settings in the tabbed area in the folder dialog box at any time prior to running the production.

After the production folders are set up, users can use the add, move, and copy-to-folder options in Review mode or the tagging window to place documents in the production folder.

As an administrator, you can search within the folder and quickly see all the documents in the folder to verify that the correct ones are in the production.

You can add documents to a production any time until the production is locked and produced. However, after the production is produced, to add documents you must first unlock the production folder, then re-lock it after the new documents are added.

Note: If the case is created in pre-10.1 release, then after upgrade to release 10.1, production folders of type Images and Mixed are affected in Imaging Tool Upgrade. For Native type production folders, all operations are available before, during and after the Imaging Tool Upgrade is performed. See the *Imaging Tool Upgrade Guide* for details.

The following options on the Mixed and Images type production folders created before in earlier releases are not available:

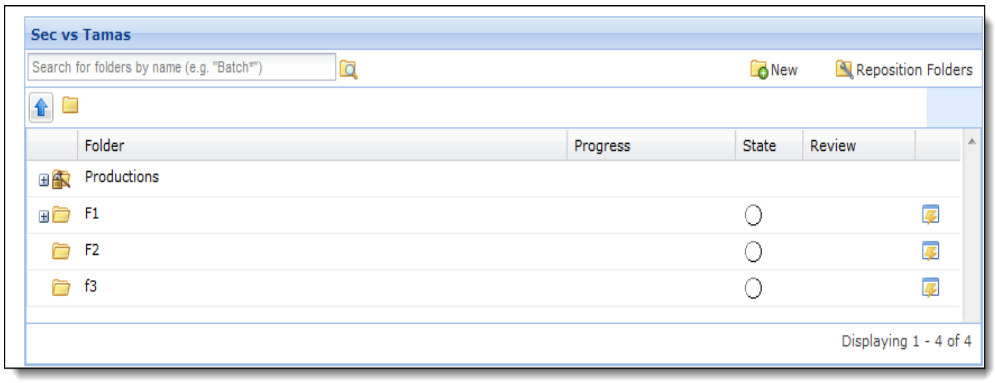
- Unlock/Unproduce locked production folders
- Lock/Produce unlocked production folders
- Edit locked/unlocked production folder

Before you begin: You must have the permission "Allow production folder management" to set production folders. The **Case Manager**, **Case Admin**, **Group Admin**, and **System Manager** roles have this permission.

After setup: Once you have setup your folders and you are ready to run a production, or perform a production export, see the Export and Production Guide

To set up production folders

1. On the top navigation bar, for a selected case, click **Analysis and Review**, and click **Folders**.



2. To add a new production folder:
   A. Select the Productions folder.
   B. Click **New** and choose the **As Subfolder of Selected Folder** option.
   C. Specify the following information:

Production Folder Settings

Field	Description
General Tab	
Name	Specify a name to identify the production.
Description	Specify an optional description.
Production Type	Specify whether to produce all documents as image files, in their native format, or a combination of images and native files depending on the file type. **Considerations:** • Natives will be produced as full document families without redactions, potentially exposing privileged information. • Selecting Native brings in complete document families.
Number of Retries	Allows you to configure the number of times to retry producing an item that has timed out. The system retries the production the specified number of times before generating a slip sheet. You can use the default value of 1 if you are running a small production to quickly validate how the production options will appear.
Redaction Set	Specify the redaction set to use for the production. See *"Setting Up Redaction Sets" on page 138*.
Excluded items: Generate Slip-Sheets for excluded items	Starting with 9.0.1, you can choose whether to include or exclude a slip sheet for any excluded, non-relevant family items. Select this check box to generate a slip sheet for the items of the document family that are not a part of the production instead of completely skipping them.
Date Produced	This field is system generated and indicates the date that the production was produced. The field is blank if the production has not been produced.

Production Folder Settings (Continued)

Field	Description
Status	This field is system generated and indicates the current status of the production set.
Embedded Objects	Choose whether to produce embedded objects separately. **Note:** Native Productions will include all items in a family. Toggle this option will extract each embedding, counting and numbering them as an individual items.
Sort Options	
Sort Production by	Choose whether to sort the productions by custodian, sent/modified date, or document ID.
Numbering Tab	
Prefix	Specify the document numbering for the production. The settings on this tab associate a production number with a corresponding document (if native) or page (if image) when the production is run. If desired, include a delimiter at the end.
Minimum number of digits	Specify the minimum number of digits for numbering the documents. The number is padded with zeros, if needed to match the minimum.
Starting number	Specify the starting number for the numbered list of documents. **Note:** The product ensures that the same production number (combination of prefix and number) is not used multiple times on the same case. If the number you specify is below the minimum allowed number for that prefix, the next valid number is displayed.
Suffix	Specify a suffix for the numbered list with a delimiter, if desired.
Sample	Shows the specified format. The sample is updated as you add numbering criteria. For example, the sample ITEM-0000012-001 shows the prefix ITEM-, numbering that includes 7 digits starting with the number 12, and a suffix of -001.

Production Folder Settings (Continued)

Field	Description
Header/Footer Tab	
Header	Choose the information to present for the left, center, and right headers (the same options are available for the footer). You can select from the following items: • None (no entry) • Custodian • Production number • Date produced • Document ID • Free text • Filename • Page number • Page X of Y **Note:** By default, the Times New Roman font is used. An administrator can configure the font family to specify header and footer text. A multi-line header/footer can be set for free text values. **Note:** In case of upgrading production folder from pre-10.1 release having Tag option value as one of the Header/Footer, then after the upgrade, the Tag option in header/footer is ignored. It will not be used for the header/footer on the newly exported images during production export.
Watermark Text	Type the word or words you want to appear as a watermark on the pages of the production. **Note:** By default, the Times New Roman font is used. An administrator can configure the font family to specify watermark text. **Note:** Do not use CJK characters in the watermark text; else, the production view fails.
Footer	Choose the information to present for the left, center, and right footers.
Font Size	Select the font size for the header and footer text. If you include a watermark, it will be auto-sized to fit across the page.

Production Folder Settings (Continued)

Field	Description
Imaging	
Maximum Time Per Item	Select imaging options for Native Imaging documents to control and improve production processing. For example, you can tell the system how to handle the imaging of large documents which can speed up production processing time. A calculator function allows you to see how your criteria would affect production processing. Specify the maximum time you want the system to spend imaging the item. While attempting to image an item, if the system exceeds this amount of time the system will stop imaging the item and move on to the next one. The default is set to 3 minutes.
Slipsheet by Maximum Limits: (Only loose files/ attachments)	Items that exceed these maximum limits produce a slip sheet and are not imaged. This only affects loose files and attachments. Email messages that exceed the maximum limits are imaged normally. This only affects loose files and attachments. Email messages that exceed the maximum limits will be imaged normally. Specify values for any of the following items: • **Page Count—** By default, page count is disabled. • **File Size—**Items that exceed the estimated file size are skipped.
Calculate Imaging	Click calculator to see what is going to be imaged based on your entries for the settings and criteria. The totals are displayed in the Results section. **Note:** These counts will always include embeddings, even if you choose not to produce them.

Production Folder Settings (Continued)

Field	Description
Slip Sheet	
Customizable Slip-Sheet Text	This option allows you to customize slip sheets with various fields which can assist in identifying exceptions. Each slip sheet will receive a bates number and the specified text will be printed in the center of the slip sheet (which is created for all items that are not imaged). The maximum length allowed for this field is 1,024 characters. You can use the following replacement case-sensitive macros: • %DocID - Document ID • %FileName - File name of the document • %FileExtension - File extension of the document • %BatesStart - Starting Bates number • %BatesEnd - Ending Bates number When produced, a slip sheet is a placeholder for any item not rendered for one of the reasons below. Slip sheet reasons: • Fully Redacted - Item was Redacted completely • Imaging Error - Unable to create image of item during production • Conversion Error - Image failed to convert to TIFF • Native Not Imported - Load File did not import Native • Native Placeholder - Bates stamped Native placeholder • Not In Production - Item was not included in production, but was part of a family where one or more items were included in production. You can apply your customized slip sheet settings at a case level by setting the property: `esa.imaging.default.slipsheettext` in the property browser. Refer to *"Using the Support Features" in the System Administration Guide* for how to use the property browser to set the property. **Notes:** • If the case property is not set, the default is: 'Image not available for this document, ID: %DocID". • Any change made to the slip sheet case-level property only impacts folders created after the case-level property is changed. Folders created prior to the case-level property change retain their previous value (regardless of if they are locked or unlocked). • If you create a case and never set the case-level slip sheet property, it will always default to the system-level default setting. This also means that if the system default changes, new folders will use the system slip sheet property (because no case slip sheet property has been set).

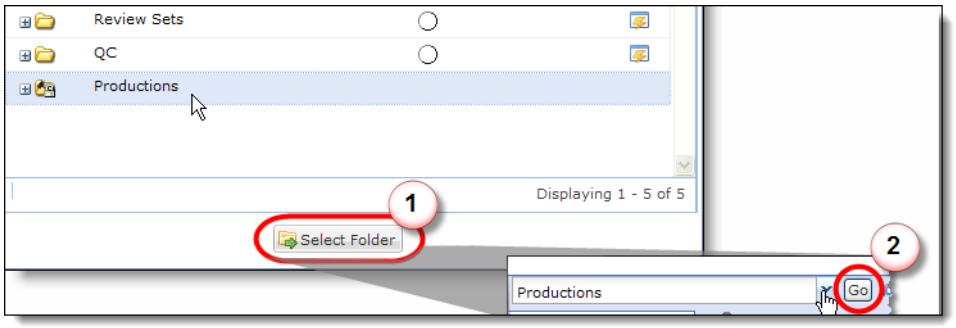
Production Folder Settings (Continued)

Field	Description
Restricted Codes	Slip-Sheets that are not able to be produced with potentially secure information will have a smaller set of Variable Codes. Slip-Sheet Reasons: • Not Produced Item - Item not produced with family Variables Codes are case-sensitive and will be replaced with: • %DocID - Document ID • %BatesStart - Starting Bates number • %BatesEnd - Ending Bates number
Results Tab	When production is complete, information about the production is displayed in this tab.

D. Click **OK** to add the new folder.

Run a Search on One or More Productions Folders

You can search for items or documents in all production folders defined in a case, or select a specific "Productions" folder in which to search.



To search a specific production folder, highlight the folder to select the subfolder, and click **Select Folder**. From the main **Analysis and Review** module screen, click the drop-down menu and select **Productions**, then click **Go**.

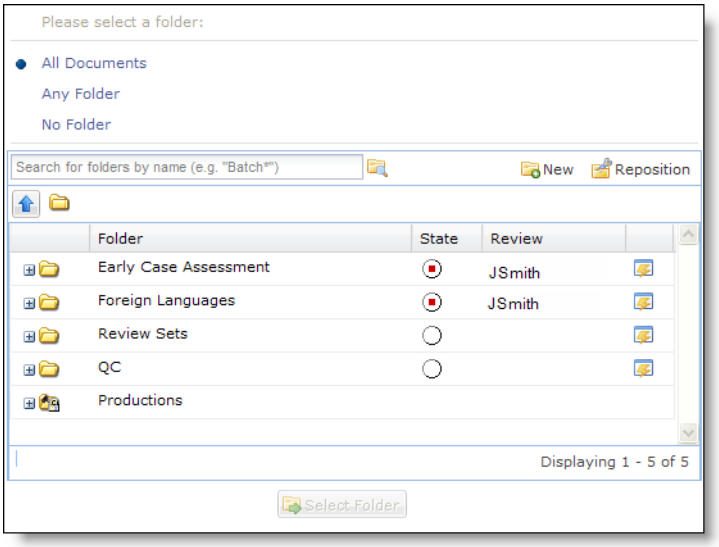
Managing Reviews

You can view and manage reviews from the **Analysis and Review** module. While reviewers work on a review set, it does not prohibit others from searching on, viewing, or tagging the documents in the folder.

To view and manage review sets from the Analysis and Review module

1. From the **Analysis and Review** module, click the drop-down menu showing "All Documents" (by default).

   The folder management pop-up window appears, with a table below showing the folder hierarchy.



   Note that as reviewers begin their review set, they can click the Status circle to let others know the current state and who is working on a particular batch.

2. Select from **All Documents**, **Any Folder**, or **No Folder** for review.

3. Click the + next to any folder (those containing subfolders) to view its contents. As you click through various levels of the folder hierarchy, breadcrumb navigation appears above the list of folders to indicate your current folder location. To return to this home screen, click the arrow icon.

4. To begin reviewing the documents in a folder, reviewers click the circle under "State". (A red circle indicates that the folder is being reviewed.) Alternatively, click edit/review icon to open the action menu, then select **Begin Review.**

   The reviewer's name appears in the "Review" column. (You may need to refresh your screen to see the red dot appear.) You can also rollover the State icon to view a tooltip message describing the current status.

Edit a Review Set

To edit or delete folders

1. From the **Analysis and Review** module, click the Folders drop-down menu.

   This menu is next to the search field and displays “All Documents” by default.

2. Click to navigate to, or select the review set that you want to stop or complete reviewing.

3. Click next to the folder you want to edit (change the name or description), then from the Actions menu select **Edit**.

4. On the Edit Folder window, change the name of the folder and/or description.

5. Click **OK**.

Setting Up Redaction Sets

Redaction is the process whereby portions of documents are concealed to protect sources or limit information on a need-to-know basis. Redaction sets allow you to apply redactions to documents in search results.

Note: Redaction sets created in pre-10.0 releases are marked as read-only in release 10.0 and are highlighted on the user interface. You cannot edit such redaction sets. You can still view the preset reason codes of these redaction sets in the Edit window and export the redaction set to CSV as well.

For information on redaction operations, refer to the *User Guide*.

Free Text and Preset Reason Codes

There are two ways to apply reason codes to redaction sets: free text and preset reason code. The free text option allows your reviewers to enter reason codes and the preset reason code option allows you, the administrator, to set up and control reason code choices. Using preset reason codes can make it easier to adhere to standards that your organization follows or that comply with regulatory mandates (for example, Privacy Information, Confidential, Privileged).

The system comes with Default redaction set, which is a free text redaction set. However, you can create additional free text or preset redactions sets if you want to apply different redaction criteria to the same set of documents. The redaction sets are available for selecting in Review mode after a user runs a search.

Preset Reason Code Considerations

- Once added, preset reason codes cannot be modified or deleted.
- Duplicate reason codes are not allowed in a redaction set using preset reason codes.
- Redaction sets created prior to version 8.3_CHF1 are handled as free-text redaction sets. They cannot be converted to preset reason code redaction sets.
- The number of Preset Redaction Reason Codes per redaction set is 50.
- A preset redaction reason code can be up to 40 characters

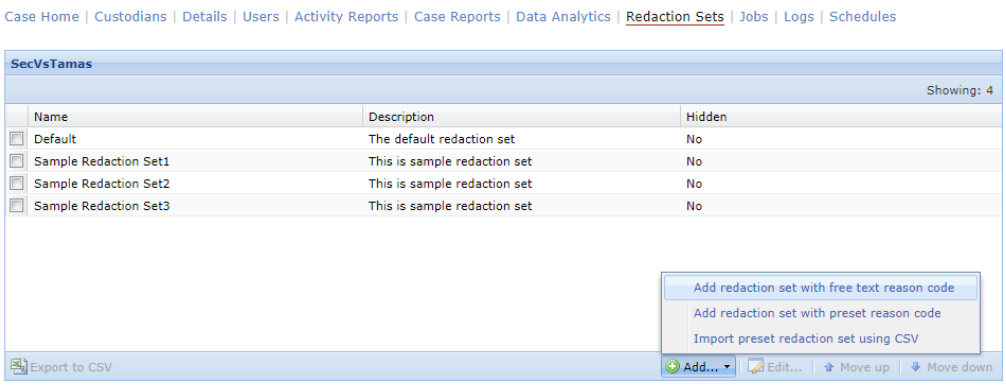
Note: Only users with the **Case Admin**, **Group Admin**, and **System Manager** roles can set up redaction sets.

To export redaction sets to a CSV file

1. On the top navigation bar, for a selected case, click **Case Home > Redaction Sets.** A list of redaction sets is displayed.
2. Select the redaction set and click **Export to CSV**.

To set up redaction sets

1. On the top navigation bar, for a selected case, click **Case Home > Redaction Sets**.



The list shows each redaction set and indicates if it is hidden.

2. To create a new redaction set:

Note: This option must be chosen at the time of redaction set creation, and cannot be changed later on.

A. Click **Add.**

B. Select from:

– **Add redaction set with free text reason code**

– **Add redaction set with preset reason code**

– **Import preset redaction set using CSV** - see step 3 if you choose this option.

Note: Free text reason codes allow reviewers to enter reason codes. Preset reason codes allow the administrator to control which codes are applied, for consistent tag searches using the codes.

In this example, **Add redaction set with preset reason code** is selected.

C. Enter a name for the redaction set.

Note: By default, the maximum number of characters allowed in redaction set name are 35. You can configure this limit by using the ***esa.case.redaction.set.name.maxlength*** property. The name field supports only the following characters—letters, numbers, white spaces, hyphens, underscores, and periods.

D. Enter description for the redaction set.

Note: By default, the maximum number of characters allowed in redaction set description are 255. You can configure this limit by using the ***esa.case.redaction.set.description.maxlength*** property. The description field supports only the following characters—letters, numbers, white spaces, hyphens, underscores, and periods.

E. Choose whether the redaction set is visible or hidden. By default, the **hidden** box is unchecked. Visibility can be edited later.

Note: Saving a redaction set with preset reason codes requires at least one reason code.

F. Click **OK** to add the new set to the list.

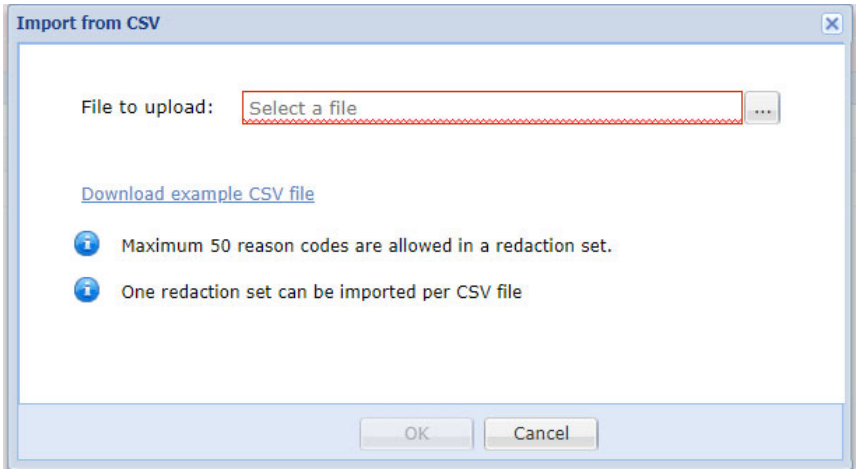
G. Select the redaction set and click **Edit** to change the redaction set name or description, add reason codes, or toggle hidden status.

Note: The name and the description field supports only the following characters—letters, numbers, white spaces, hyphens, underscores, and periods.

3. To import preset redaction set using a CSV file:

A. Click **Add**.

B. Select **Import preset redaction set using CSV**. The Import From CVS dialog appears.



C. Browse and select the CSV file to upload.

Note: Only one redaction set can be imported per CSV file, and that redaction set can only contain a maximum of 50 reason codes, by default. The number of reason codes per redaction set can be configured using a property ***esa.case.redaction.preset.reasons.maxallowed***. The default value is 50.

You can use the property ***esa.case.redaction.set.maxallowed*** to define the maximum number of redaction sets to be added or imported. By default, the value is set to 20.

If you have more than 50 reason codes, you can use the Import option multiple times. It is recommended to be careful while importing or adding a redaction set as these cannot be deleted once added or imported.

The CSV file must be in a specific format. You can see an example of CSV file by clicking Download example CSV file.

The name and the description field of the preset redaction set (imported using CSV file) supports only the following characters—letters, numbers, white spaces, hyphens, underscores, and periods.

D. Click **OK**.

4. To change the placement of the redaction set, select the set and click **Move Up** or **Move Down.**

Viewing Case Participants and Groups

For information about how to view groups and participants for a single case, refer to the following topics:

- *"Viewing Case Participants" in the next section*
- *"Viewing Groups" on page 143*

Viewing Case Participants

You can view the name, email address, and group name of all internal and external email addresses discovered in a case. All external participants belong to the External group.

You can also define arbitrary groups of participants to find documents sent or received by any member of the group (see *"Viewing Groups" on page 143*).

To view participants

1. On the top navigation bar, for a selected case, click **Processing > Participants**.
2. To show the non-primary display names/email addresses associated with a participant, click the arrow to the left of the name. Click the arrow again to hide the names.
3. To show the non-primary display names/email addresses for all participants in the list, click **Expand All**. To hide the non-primary display names/email addresses for all participants, click **Collapse All**.

Note: The number in parentheses to the right of the primary email address indicates the total number of primary and non-primary addresses for that participant.

4. To search the list of participants:
   A. From the **in** menu, select the column to be searched (**Full Name**, **Primary Email Address**, or **Group**).
   B. Enter the first few characters of the search text in the **Search for** field (use a "*" to indicate any text, such as "*ob*").
5. To view additional participant details (if any), click the participant name. Note that internal participants may have multiple addresses under the same name. Click **Back** when finished.

Viewing Groups

Internal email addresses (internal participants) are automatically grouped into groups based on the department data retrieved from your Active Directory server (if any).

Note: Two additional groups are shown only in the search results filter. The External group includes all external users; the Internal group includes internal users who do not belong to any other group.

To view groups

1. To view just the manually defined groups or the groups found in the Active Directory, select **Manual** or **AD Department** from the **Type** menu.

   To search for participants in a case, use Advanced Search in the **Analysis and Review** module. For more information, refer to the *Veritas eDiscovery Platform User's Guide*.

Managing Batches

Batches are used to group sets of documents that are processed into the system or exported/printed out of the system. Batches allow you to manage the flow of documents through the system and maintain a detailed audit trail of how a document enters and leaves the product.

The Manage Batches screen summarizes the batches that have been created for the case, and provides detailed information about the batch. It also allows you to change the label of one or more batches. You can also optionally create batch labels for export and print jobs.

To manage batches

1. On the top navigation bar, for a selected case, click **Processing > Batches**.

   The list shows each batch and includes the label, type of job, user that ordered the job, start and end time, and duration.

2. To change the label on a batch, click **Relabel**, enter the new label, and click **OK**.

Pre-Processing Navigation

For information about pre-processing your data, refer to the following steps:

Overview

Pre-processing your data can reduce cost by enabling you to target specific data within a source collection.

Note: Note: You must have a licensed, installed Pre-Processing module, and pre-processing enabled on your system at the time of case setup to use the following features.

The pre-processing module offers:

- Pre-Processing Analytics

  Visually 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 eDiscovery budgets and timelines.

  Note: Pre-processing analytics are always available for loose files (with or without the module license). However, analytics for data in PST and NSF files are available only with the license and the case settings option for collect pre-processing analytics.

- Advanced Pre-Processing Filters

  Enable users to interactively filter data by custodian, date, strong file type, and file size prior to processing. The product's one-click filtering of custom file and "NIST list" items can significantly reduce downstream processing and review costs. This also enables pre-processing of LEF, and E01 files.

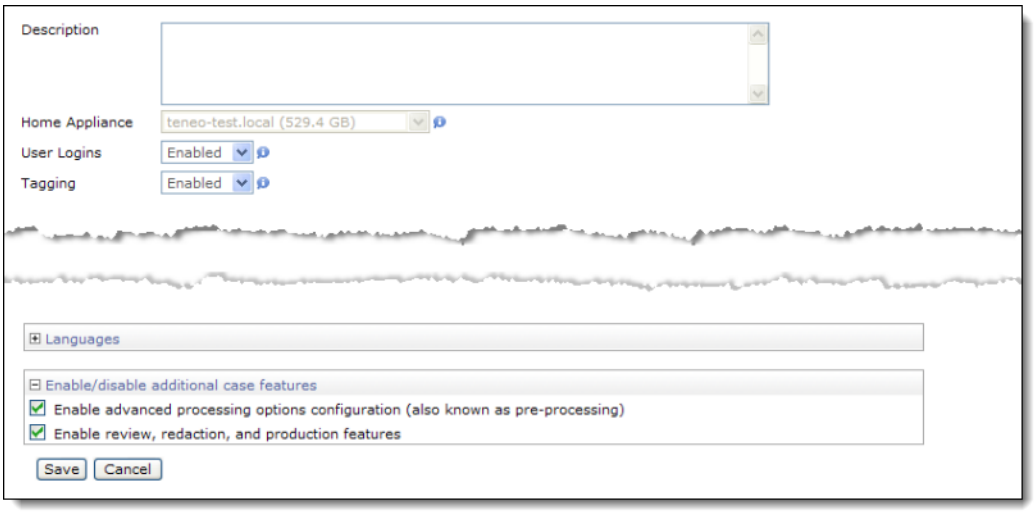
Step 1: Enable Pre-Processing

You must have licensed the Pre-Processing module in order to discover and process LEF files, and to get pre-processing analytics for PST and NSF files.

How to enable pre-processing

Enable the Pre-Processing module when you create a new case.

From the **Settings** page, select the **Enabled advanced processing options configuration** (also known as pre-processing) option.



Step 2: (Optional) Exclude Documents on the NIST List

Collected document sets often contain many files which are not user generated but are instead common application programs, help files, or DLLs. These files are identified and tracked by the National Software Reference Library and commonly known as the "NIST List". In addition, most companies have standard software images that they use for creating desktops, laptops, and servers, and often want to exclude those files from eDiscovery since they are not relevant.

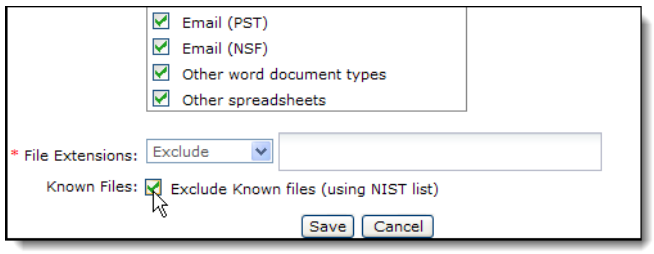
The product provides the ability for users to exclude both the standard NSRL (NIST list) files and to upload company-specific file hash lists into the system. Once the known file lists have been set up by the administrator, matching documents can be excluded with a single click while setting up a document source.

Note: Excluding documents on the NIST list occurs when you add a source.

How to exclude files on the NIST list from processing

1. From the **Processing** module, for the selected case, click **Sources and Pre-Processing**.
2. Select **Add Case Folder Source** and click **Go**.

   The Add Case Folder Source page displays.
3. Select the **Exclude Known Files (using NIST list)** option.



   Note: If you select to Include all file extensions, in the text field following your selection, you must specify which file extensions to include. Otherwise, no files will be processed.
4. Click **Save**.

Step 3: (Optional) Merge Custodians

When you have two or more same, or similar custodians (representing the same custodian) you can merge them into one unique custodian assignment. This is optional, and can be done either before or after processing your case data.

Note: If you process your case data first then merge custodians, you must rerun post-processing to update changes in custodian assignments in the case.

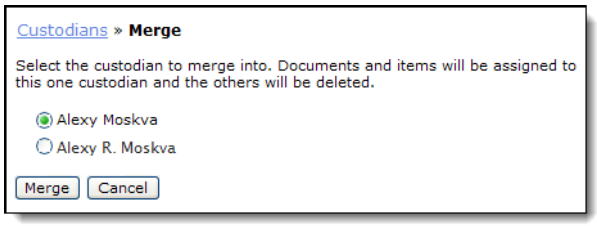
How to merge custodians

1. From the **Processing** module, for the selected case, click **Custodians**.

   Your list of custodians displays.

2. Select two or more custodians that you want to merge into the same custodian assignment.

3. Click **Merge**.



   Note: The single custodian you select will automatically be associated with all documents and items previously associated with both. All other related custodians listed on the Merge page will be deleted.

4. Click **Merge** to confirm the single custodian assignment.

   For further details about Custodian Merge (or to un-merge custodians) refer to *"Managing Case Sources and Custodians" on page 49*.

Step 4: Analyze and Filter Sources

Analyzing and filtering sources is often an iterative process. You perform an initial analysis and then begin adjusting the filters and the set of sources that encapsulate the case data you need to process.

To access the Pre-Processing Interface

1. From the **Processing** module, for the selected case, click **Sources and Pre-Processing**.
2. Click the **Pre-Processing Options** tab.

   Note: The Pre-Processing Options tab is only available if you have licensed Pre-Processing and have enabled it for the case during case creation. (This tab is not available when data in a case is currently processing.)

   The Pre-processing interface, or Pre-Processing Options screen, displays.

Understanding the Pre-Processing Interface

Analyzing and filtering a collection source consists of the following tasks:

1. Selecting sources.
2. Choosing filters.
3. Analyzing how the filters impact the type and amount of data to process by manipulating the visualization options.

   Tip: Reduce the filter pane to optimize the Pre-processing view.

About Pre-Processing Analytics

The right pane of the pre-processing interface is the visual display pane. It enables you to visually summarize overall document set characteristics and presents detailed analysis by custodian, timeline, and file type. You can quickly confirm that all case data has been collected and uses the detailed statistics to accurately estimate eDiscovery budget and timelines.

You can view all four visualizations in table form by using the View as drop-down and export the data in CSV or XML format by using the Export drop-down. The Data Axis drop-down applies only to the chart view and can be set to either Case Size or Relative. Selecting Case Size sets the maximum value on the data axis just above the case size, while selecting Relative sets it to a value near the top value in the current chart. The charts display processed, unprocessed, and "selected to process" data in different colors.

Analysis Options

The visual display pane provides four different analyses of your case data. Viewing them helps to quickly confirm that all case data has been collected and provides accurate estimates of downstream eDiscovery volume or effort.

1. **Summary:** This waterfall chart provides a summary view of your data and is an easy way to view the amount of data you need to process after excluding known files and applying processing options.

   Note: If errors occur during pre-processing, a red bar appears indicating the total number of files with errors. This can include any corrupt, password-protected, or unrecognized files.

2. **Document type:** This bar chart displays the distribution of your data by document type such as PST or NSF emails, office documents, or multimedia, for example.

   Note: You must have the Pre-Processing option turned on, and the module licensed and enabled during discovery on the source. Otherwise, files are included/excluded based on the Modification date.

3. **Custodian:** This bar chart provides a representation of processed, unprocessed and selected to process data by custodian.

4. **Timeline**: This bar chart displays the distribution of data across a timeline. You can choose to aggregate the data by day, month, quarter, or year. You can also choose all dates, or only dates for which data is present.

   Note: Items without a sent date, shown as (none) in pre-processing timeline charts will continue to be shown as unprocessed (not selected) even if they are processed. This can occur even when you have not applied any date filters to the source.

   If you do not specify a date filter for a source or sub-source, then items with no date will be shown as selected to be processed in other charts and in Manage Sources. However, be aware that data will not, in fact be processed and will be reflected in timeline charts.

Using Pre-Processing Filters

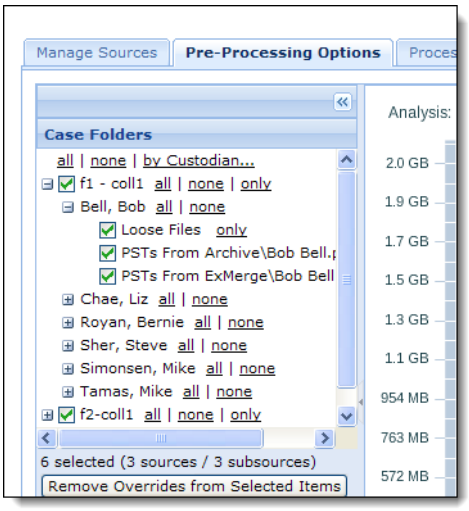
The Processing module provides the ability to choose case folders at different levels of granularity, and then apply processing options to them. The Case Folders and Pre-Processing Options for Selected Items panes are used together for this purpose.

You can interactively filter data by custodian, date, strong file type, and file size prior to processing.

Step A: Select the Source

Case folders can be selected from the Case Folders pane in a variety of ways. You can select entire collections, choose by custodian, or select specific files (PSTs/NSFs can be selected individually, while loose files are grouped together).

- See "To select case folders by custodian"
- See "To select specific custodians"

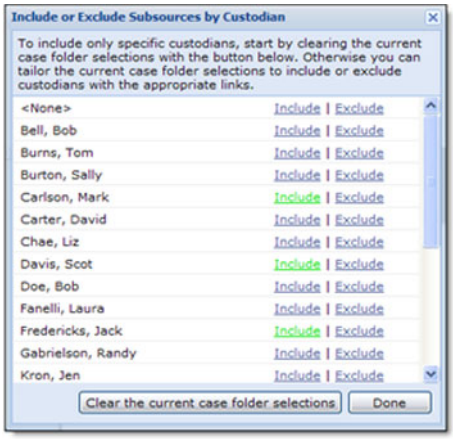


Note: Including or excluding PST and NSF files is based on Sent time.

To select case folders by custodian

- Click the **by Custodian** link in the Case Folders pane.

  The **Include or Exclude Subsources by Custodian** window appears.



To select specific custodians

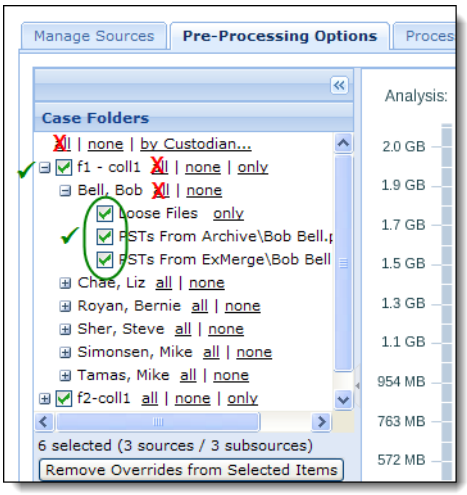
- Click **Clear the current case folder selections** (as shown in step 1) and select specific custodians by clicking **Include** next to each name. Selected custodians are highlighted in green. Items for which no custodians have been identified are represented by a custodian named **<None>**. When you are done selecting custodians, click **Done**.

Step B: (Optional) Discard Overrides for Selected Items

If processing options have not been specifically set for a folder, the folder inherits the processing options of the case folder source by default. To change these options, you can select the folder and apply specific processing options (discussed in the next section). This will override the parent level processing options the folder had previously assumed, and an exclamation mark (!) will be displayed in front of the folder name.

If at some point you want to clear the processing options set for this folder, select the folder and click **Remove Overrides from Selected Items**. When clearing processing options assigned at the folder level, the folder inherits the processing options of its parent folder.

CAUTION: Selecting "all" rather than just the case folder source alone may result in overrides on every sub-source. Instead, to avoid overrides, select the sub-source, and the individual source file types. In this example, the specific collections are selected.



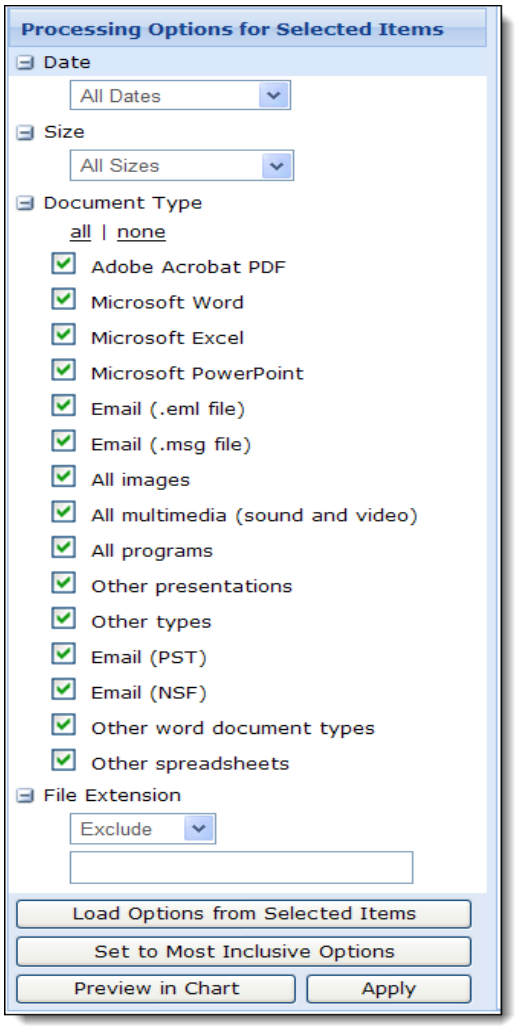
Step C: Apply Processing Options to Selected Items

Processing options allow you to specify content for processing within files and folders you have chosen in the Case Folders pane. It allows you to easily filter data by date, file size, document type, and file extension and narrow down the documents to be processed.

How to view current processing options.

To view the processing options for a folder or a set of folders, in the Case Folders pane, click **Populate Values from Selected Items**.

The processing options pane is populated with the current options for the selected items. If a specific value is not the same for all the selected folders, the option for this value will be unavailable. (Clicking the "grayed out" checkbox allows you to modify it.)



How to set the widest range of options

If you have previously set processing options for a folder or set of folders and want to discard them and select everything for processing, click **Set to Most Inclusive Values**. This resets the options to the widest ranging values. Specifically, clicking this option selects all the Document Type checkboxes, sets Date to all dates, Size to all sizes, and clears the **include/exclude File Extension** checkbox.

How to view the effect of current processing options

Click **Preview**. The current settings are reflected in the visual display pane.

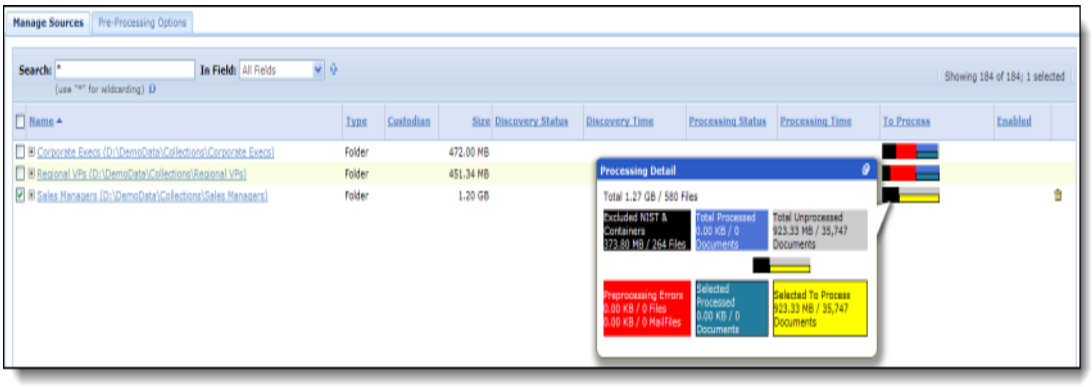
How to save the processing options:

When you are satisfied with the results of the processing options in the preview, click **Apply** to save changes for the current folder selection.

Step 5: Verify your Saved Processing Details

After iterating through all your folders and saving the processing options for each folder, click the **Manage Sources** tab to go back to the summary page.

Detailed reporting is provided on this page to indicate the total number and volume of items excluded due to known file filtering, and shows the total volume and number of items to be processed based on user-specified filtering criteria.



Run and View Reports to Prepare for Processing

Most of the reports can be used before and after processing. For each report type, you can select the batch label and run a report for that specific batch. To access the reports, go to the **Processing** module for a selected case, and click **Reports** to select the report you want to run.

Note: Batch level reports are available for processing batches in cases created prior to version 6.6. For cases created in 6.6 and later, additional reports for discovery batches and processing batches are also available.

Before Processing…

- View file errors found during discovery.

  Run the **Discovery Errors** report before processing to view an initial list of errors found during discovery, including PST and NSF files that failed discovery. Before processing, this report provides a list of issues that you can address without processing the entire batch first. When selected, you can also choose from a list of Reason Codes to search for a specific error type. (After processing, run this report again for a comprehensive list of file errors. See Step 7.)

- To view all items not yet processed (including items selected or excluded based on your pre-processing options), run the **Not Processed Documents** report.

  Note: This report is available for loose files only, and does not include data from PST and NSF files. To view other file types in your case, run the **Not Processed Documents** report after discovery has completed, and check the Type and Extension option.

- View files excluded by the NIST list

  To determine which files are excluded from processing due to the selected NIST list, run the **Not Processed Documents** report**.**

- View the filtering options selected for the data set.

  Run the **Discovery and Processing Options** report to list options selected for Discovery or Processing batches of documents or Processing Source data.

- View documents containing errors after import from a load file source.

  Run the **Load File Discovery Errors** report to view a list of documents that contained errors upon importing a third party load file source in pre-processing.

  For details about importing load files, refer to the *"Load File Import Guide*".

Step 6: Start Processing

After reviewing processing details for your sources, select the sources and start processing. Statistics are tracked and monitored across multiple processing runs and are fully incremental, so users can easily process an initial set of documents for early case assessment based on the most critical document types and time periods and then easily expand out the range of documents processed as the scope of the case expands.

Step 7: Review Processing Results

To learn about the files that were processed (or did not process), generate the processing reports found on the **Processing module** > **Reports** screen.

For more information, see *"Generating Processing Reports" on page 99*

Run the following reports to view the different aspects of your case files.

After Processing…

- Review a complete and comprehensive list of locations for de-duplicated documents. During processing, de-duplication is performed across all document sources and all custodians.

  Run the **De-duplication** report for a complete and comprehensive list of locations and sources for de-duplicated documents.

- Review a list of custodian level de-duplicated documents.

  Run the **De-duplication by Custodian** report to see, at the custodian level, how much the document set has been de-duplicated during Processing.

- Review the filtering options chosen for the data set.

  Run the **Discovery and Processing Options** report to generate a list of options used.

- Review all errors encountered during processing.

  Run the **Discovery Errors** report after processing to see a comprehensive list of errors found during processing, including PST and NSF files that failed to process. Errors that can only be discovered during processing include MBOX conversion errors.

- Review a list of files that were not processed. For example, whether a document is discovered but not processed, or if it is excluded from Discovery or Processing or both based on the different exclusion criteria (such as date range, size range, file types and extensions, de-NIST list, and container file count).

  Run the **Not Processed Documents** report to display files that were not processed or excluded by the case's pre-processing options or exclusion criteria.

  Note: This report is only available after attempting to process the data.

- Review the documents that produced Discovery errors during Load File Import.

  Run the **Load File Discovery Errors** report to identify Discovery errors affecting Load File Import.

  Lists the documents that were included and imported from a load file source. For more information, refer to the *Load File Import Guide*. (Shows loose files only).

- Review "Other Type" documents (these are documents that do not fit into any other of the categories listed under Document Type in **Processing > Sources and Pre-Processing > Pre-Processing Options**).

  Run the **Other Type – Extensions** report for a list of document that are not listed according to their type.

- Review a summary of all processed files within the case.

  Run the **Processed Documents** report for a complete list of files processed including any errors that were encountered.

- Review and track data during Discovery and Processing phases for reconciliation.

  Run the **Processing Reconciliation** report to get summary and detailed reports for the list of files and the associated files on disk and document/attachment counts needed for reconciliation. This report identifies which files get discovered, extracted, excluded, and the exact count of documents and attachments that are processed or dropped from those files.

Processing Exceptions

For information about processing exceptions, refer to the following topics:

Introduction

With electronically stored information (ESI), the question is not really if you will run into processing exceptions, but when. The product's transparency is not just present in the search and analysis portion of the solution, it is a fundamental basis for every function and feature within the product. Through audit logs and robust reporting features, you get a complete end-to-end accounting of a customer's eDiscovery workflow from pre-processing through production.

Why do Exceptions Occur?

When processing large amounts of ESI that, until collection, has been largely unmonitored and unstructured, it is quite likely that some files are in a corrupted state, password protected, or encrypted. When the product encounters these kinds of file states, they are identifies them as potential issues and logs them as exceptions.

While over 400 file types are supported, there is a chance that in large data sets that some files will be encountered that are not currently supported. If ESI is encountered during processing that falls outside of this set of supported file types, an exception is logged as well.

What are Exceptions?

Exceptions are typically documents that are flagged with a warning and processed. These documents contain some aspect of processing that either did not perform as expected or contain an attribute that warrants additional consideration. For example, a PDF file that has been password-protected would be flagged to alert administrators of this file. Flagging the document enables the administrator to make informed decisions regarding remediation as well as report on these exceptions to the legal team if necessary.

How Are Exceptions Handled?

The product logs contain every exception encountered during processing. These exceptions are retrievable using exportable detailed reports, filters in the **Analysis and Review** module, and advanced searches.

There is also an **Exceptions** screen located in the **Processing** module of the application that allows an administrator to browse exceptions across an entire case, and even filter those results by batch, custodian, etc. The administrator can then run, and export customized reports for all exceptions, all warnings, specific exceptions, or specific warnings. These reports contain all of the information that would be necessary for an administrator to perform a remediation workflow on the files or determine whether the files must be analyzed outside of the application.

In some cases, these files will be damaged or corrupted to such a degree that third party tools are required to "repair" them. In these instances, it is important to preserve chain of custody and data integrity by only remediating copies of the original source data. Remediated files can then be reprocessed in order to make that content available to users for analysis and review. See *"Image Remediation" on page 88*.

Steps For Managing Exceptions

This part of the process is typically referred to as *remediation*. This is where the unprocessed Electronically Stored Information (ESI) that could not be processed is analyzed and decisions must be made on how to handle or repair these documents.

Remediation Workflow

Essentially, administrators can add fixed documents as a new source, or as a new subdirectory in the existing case folder source. At that point, administrators can then process with discovery and a new batch name.

Note: During the remediation process, the administrator should document each step to preserve the process in the event it was ever called into question. While actions performed within the application are logged and auditable, remediation steps taken outside of the product and with third-party tools are not. Documentation is a key component of defensibility.

The following workflow outlines the **Case Admin**'s tasks and available options.

- **View Exceptions**

  First, the administrator can access the Exceptions screen in the Processing module for a selected case. All exceptions encountered are displayed for administrator review to determine next steps. See *"Document-Level Errors and Warnings" on page 166* for further details.

- **Export Files**

  Next, the administrator may choose either to export the specific files as CSV, versus an actual export of documents that were affected by the exceptions. Because these files are a duplicate copy of the original source ESI, they can be handled and remediated without fear of affecting the original source ESI.

Note: After export, each file should be examined and remediated if possible (utilizing third-party tools if necessary).

- **Reprocess Files**

  After the files have been remediated to the extent possible, the Administrator can then attempt to process the files again in the application. The Administrator should name this processing batch with a descriptive name that identifies it as a "Remediation Batch" or a similar name. This ensures that administrators and users can quickly and easily return to this exact document set within the product interface if required.

This process can simply be repeated in an iterative fashion until all exceptions have been remediated.

IMPORTANT: Some files cannot be processed either because the content is not supported by the application or the file is too damaged to be remediated effectively. These exceptions will be noted in the exceptions reports, and after doing as much remediation as possible or as desired, reports detailing the remaining exceptions will be retained within the application as part of the case while it is active. These reports will also be retained as part of the backup or archive of that case.

Remediation Best Practice

Administrators may also want to perform a cost/benefit analysis during remediation. In some instances, the files that are generating exceptions occur because they are not typically analyzed during eDiscovery. They may be system-generated files that do not require review or that would be culled out early in the analysis workflow. There is no real benefit to remediating or attempting to remediate these files if they will never be reviewed. Again, documentation is key, so logging these decisions is critical to maintain a defensible workflow that can be recreated. Any files that are exported for remediation from the product will be logged, and each file that is processed again after remediation will be logged as well.

How Should Source Data Be Handled Before, During, and After Collection?

While the application does not alter the integrity of source data during processing, there are some cases where it is impossible to process certain data types without making a small modification.

For example, PST and NSF files will have their last accessed time modified when the crawler service accesses those files for processing. This is not a unique issue to the product, it is a function of the MAPI API and Notes API used to access these files. So, it is considered a best practice to preserve a pristine copy of your collected source data throughout the duration of the case. Then, a working copy can be made to process into the product. At the very least email container files such as NSF or PST files should be preserved before processing data with the application.

Also, users should never open PST or NSF files that are being processed or have been processed by the application. The application assumes that data will not be modified after it has been processed. If data is modified after processing, users could experience unpredictable behavior that may affect the integrity of searches, viewing, redaction, production, and export. Additionally, a single copy of a file should not be processed by more than one case. Files can only be processed safely in a single case. If a file needs to be processed across multiple cases, a copy should be made for each case, and then the data should be organized by case at the source location.

When moving data from one location to another, it is possible to modify metadata such as the last accessed time and last modified time. There are third-party tools such as Microsoft's Robocopy and MicroForensics' Evidence Mover that will preserve metadata and verify a file's hash value in both the source directory and the destination directory to ensure that the files were not modified when being copied. These steps should be taken when making working copies of source ESI for processing within the application if metadata integrity is of great concern in the case.

Conclusion

Exceptions are a part of eDiscovery processing, and each case has specific requirements that may dictate a greater level of remediation or attention to these exceptions. The most important step to take now, though, is to create a consistent, repeatable workflow that can be used in 99 percent of cases, and when a specific case warrants additional scrutiny or diligence, the workflow can be modified to suit that circumstance. When that workflow is modified, it should be documented. The product's exception handling and reporting are every bit as transparent and robust as the rest of the solution, and the features and tools available within the product to deal with exceptions will enable a user to provide very detailed information about what data was processed, what was not, why it was not, and whether that data was processed again.

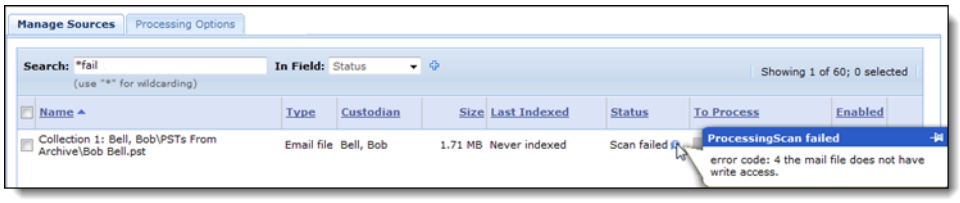
Overview

There are three levels at which processing errors and warnings occur in the product:

- Source-Level Exceptions
- Pre-Processing Errors
- Document-Level Exceptions

Source-Level Errors

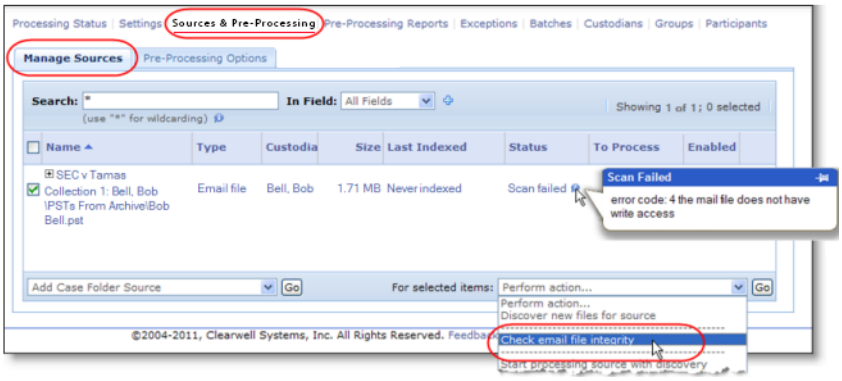
Errors at the document source or sub-source level occur when the application is not able to access a server, or when an email container file (such as a PST or NSF file) presents a high-level error when the application attempts to open it for processing. These types of errors may occur during processing or during pre-processing and are presented in the **Processing module > Sources and Pre-Processing** screen (Manage Sources tab) in the Status column. Additional detail for an error may be displayed by hovering over the infobubble in the Status column.



When searching for source-level exceptions, use wildcards with keywords. For example: *fail or *warn. Then hover over the infobubble to read the error message.

Integrity Scan Errors

The application can automatically check the integrity of newly added email files during discovery.



To do this manually, select the Check email file integrity option to perform a pre-processing scan of PST/NSF email files for potential processing issues. Any file(s) identified to have an issue will be disabled from processing. The following is a list of common integrity scan errors.

Common Integrity Scan Errors

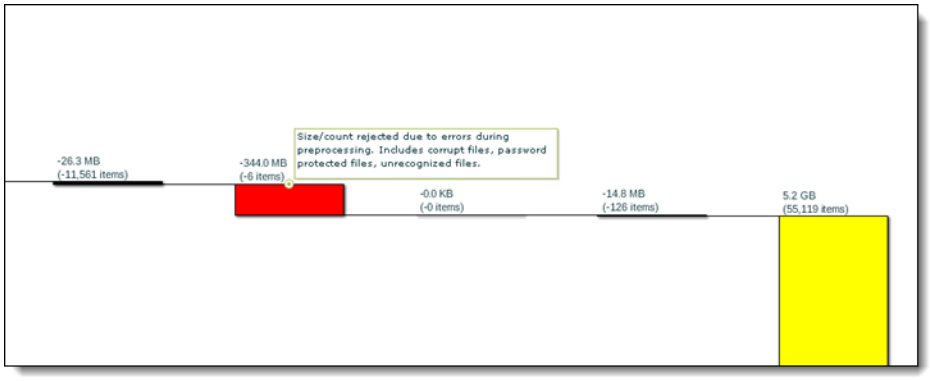
Scan Error	Explanation
Not a valid PST/NSF file	Looks at the digital signature of each email container file (PST/NSF) to determine if it is a valid email container. If not valid, the file will be marked with a "Scan Failed" status and a "Not a valid PST [or] NSF" file message in the infobubble.
Email file is read-only	Both PST and NSF files must have write access to be processed in the application. If the scan detects missing permissions, the following errors are generated: "For PST files, error code: 4 "The mail file does not have write access", or for NSF files, error code: 8 "The mail file does not have write access for current user."

Common Integrity Scan Errors

No read access to email file	Both PST and NSF files must have read access to be processed in the application. If not, the following error is generated: "The mail file does not have read access for current user."
Mail file has emails with empty recipients	During a scan, PST files may be corrupt through Exchange 2010's "pre-update Rollup 3". They are subsequently repaired using Microsoft's SCANPST tool packaged with Microsoft Office. If the application finds files with missing recipients, the following error message is generated: "The mail file has emails with empty recipients". Refer to the Release Notes for more information.

Pre-Processing Errors

Errors prior to processing documents in the application occur when a case folder source is discovered, or during discovery of new files for a specified source. This pre-processing step identifies any problematic documents (with pre-processing enabled). All other file level errors are available post-processing via reporting. This is also available in graphical table format and CSV format. (For information about load file import errors during pre-processing, refer to the *"Load File Import Guide" Load File Import Guide*.)



Be sure to run a Discovery Errors Report to list all errors found during the discovery phase (**Processing >Reports**). You can also apply a reason code for the selected batch, and choose CSV or XML format. Your results will be available as a container ZIP file.

Understanding Processing Errors

The following table shows the possible processing error messages that can occur, followed by a description of the error.

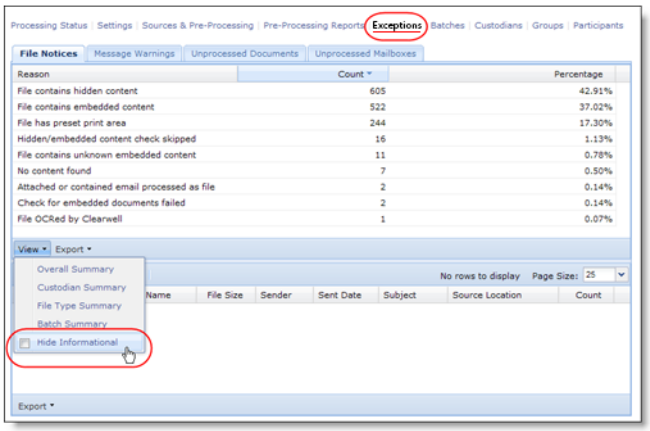
Processing Errors

Processing Error	Explanation
Error 15003 when processing PST files	Error 15003 means the PST file is not accessible. Possible reasons are that it is locked, password protected or the file share is inaccessible. More detailed information should be available in the Status field infobubble or the PSTCrawler logs.
Error 15016 when processing PST files	Error 15016 means the PST file is not accessible. Possible reasons are that it is locked, password protected or the file share is inaccessible. More detailed information should be available in the Status field infobubble or the PSTCrawler logs. This normally maps to MAPI error 0x80040600. 0x80040600 means MAPI_E_CORRUPT_STORE. This generally means that the PST file is corrupt. ScanPST can sometimes repair it. If however you are crawling over a network share, then the profile might get corrupt if there are any glitches on the network, and that will produce the same error code, even though the file isn't corrupt. So either the PST file itself is indeed corrupt, or there was some unexpected profile corruption when trying to access that file.
Other error codes	Other MAPI error codes can be referenced at the Microsoft KnowledgeBase: http://support.microsoft.com/default.aspx?scid=kb;en-us;238119

Document-Level Errors and Warnings

Document-level exceptions occur when the product is not able to process a specific document within a source/sub-source. The application tracks this information in tracking tables that are accessible to the **Case Admin** after processing is complete. This information may be exported from Processing > Exceptions.

You can view document-level errors and warnings by custodian, file type, batch, or as an overall summary. Click the **View** menu and select the summary you want to display, or view additional informational-level messages.



To view informational messages, clear the **Hide Informational** option (selected by default).

You can view Informational messages from both the File Notices and Message Warnings tabs. (File Notices are captured for both loose files and attachments.)

You can also customize your view. Click the arrow next to the Reason, Count, and Percentage columns to sort errors, as well as view errors by: Overall Summary, or Summary of Custodian, File Type, or Batch.

To export all (or individual) exceptions, click **Export** at the bottom of the screen, and choose **CSV** format for easy review.

Reporting of Document-Level Errors and Warnings

The application tracks document level processing errors and warnings and places them into downloadable reports for the **Case Admin**. For each case, processing exception reports may be found under **Processing > Exceptions**.

The 4 report categories covered in this document are:

- Unprocessed documents
- Message warnings
- File notices
- Unprocessed Mailboxes

If the **Document ID** column in any of the reports is blank, the message/file was completely dropped by the application and is not available for searching through the user interface. If it has a value, then it was at least partially processed and is available through the user interface. In certain cases, a document level infobubble is available in the user interface to display information about a warning to the end user.

Unprocessed Documents Report

The following table shows the possible unprocessed document error messages that can occur, followed by an explanation, reason code, and description of the error.

Unprocessed Document Errors

Reason	Explanation	Description Field
Language Processing Error	Error in figuring out language boundaries while processing the document.	Error processing document in language boundary analysis
Processing Error - An unexpected error was encountered during indexing.	Error during email duplicate elimination using email locator service.	com.teneo.esa.indexer.duplicateeliminator.DuplicateEliminatorException: [#20054] Error during email duplicate elimination using email locator service
	Error resolving participant for the email address.	DESC=Error resolving participants reason - com.teneo.esa.indexer.participants.ParticipantResolverException: [#20062] Error resolving participant for the email address
	Unable to find Neaftid. [No email attachment file to (document) ID.]	
Crawler dropped	The content of the document could not be retrieved for some reason, generally due to a MAPI error.	[#10024] Message referred by a URN could not be retrieved There are also a number of other MAPI-type errors that may fall into this category.
	The message is encrypted so it could not be read. Lotus Notes only.	Encrypted Message
	Path contains supplementary Unicode characters. The application does not support supplementary Unicode characters in the path and filename. (Loose file and container.)	
	PKI decode failed; error decrypting message (due to lack of certificates).	

Unprocessed Document Errors

Ignored document type	Administrators can configure the product to ignore certain document types (for example, calendar entries or contacts).	EsaMapiSearchFolder: ifolder->OpenEntry() failed EsaMapiSearchFolder: ifolder->OpenEntry() failedEXEX
Duplicate processing error	When the application went to index the email, it found that a duplicate for it was already in the process of being indexed. The application then waited for indexing to complete on the original message before marking the duplicate as successfully indexed. However, after retrying a number of times, The application was unable to get a definite status back about the state of the original message. As a result, one of two things happened: Even though the timeout limit was reached, the original was still successfully indexed. In this case, the duplicate location is lost even though the original content was ultimately indexed successfully. The original was not successfully indexed. In this case, both the original and the duplicate will not be searchable.	Unable to find original document
Container processing error	Error extracting contained file names from the container.	
Password protected container		
Container with large number of files	This error occurs if the combined file size of items within a container file exceeds 10K.	

Message Warnings Report

Messages can be flagged with multiple warnings and as a result may appear in multiple categories.

Message Warnings

Reason	Explanation	ERROR or Informational Message
Error Processing Attachment	A loose file or message attachment was processed with an error. (See File Notices for details.)	ERROR
Sent time modified	The sent time of the message was modified because it was missing or invalid.	Informational
Sent time missing	The sent time of the message was missing or invalid.	Informational
Sender missing	The sender of the message was missing.	Informational
Sender modified	The sender of the message may have been replaced with a modified name.	Informational
Crawler truncated	Email content truncated by crawler. The length of the email body exceeded. The application's maximum body limit of 512K was reached, so it was truncated. Any message content prior to the 512K limit was indexed normally.	Informational
Message partially indexed	The email was partially indexed because its content exceeded the maximum token limit of 100,000 (v3.0) or 500,000 (v4.0).	ERROR
MIME encoding truncated	Document contains mime encoding content. The application detected that the message body contained MIME encoded content that it was not able to process. A document for the email was still created in the index and searchable content was indexed to the best of the product's ability.	ERROR

Message Warnings

Notes truncated	It is a NSF notes document and the application found it was truncated.	Informational
Signed Message	The message was signed by a PKI certificate.	ERROR
Encrypted Message	The message was encrypted by a PKI certificate.	ERROR
Contains embedded images	This message is typically seen with NSF files, when the message contained embedded images.	Informational
Attachment/file information flagged	A loose file or attachment to a message was processed with an informational file flag. (See File Notices for details.)	Informational

File Notices Reports

The File Notices table contains warnings about issues that were encountered during processing. These issues were not severe enough to cause the document to not be processed, so they are flagged for review.

In the case of file notices, a "stub" is always created in the index (along with a document/attachment ID) because at a minimum the application had enough information from the crawler to create the basic information about the document in the index.

Documents with file flags are marked with an infobubble in the UI.

File Notices

Reason	Explanation/Notice Details	ERROR or Informational Message
No content found	The file was not empty, but no text content was found in the document to index (for example a .gif image) This notice may contain one or more of the following details: • No filter available for this file type (0x0004) • Could not create attachment file • <blank> • Rendering of this format is not supported (0x0033) • Converted text file does not exist	Informational

File Notices

File is empty	The file is zero bytes long or otherwise found to be empty. This notice may contain one or more of the following details: • Could not create embedded message as an attachment, SIZE=0 • Attachment filename is null or empty • File is empty (0x000A) • Unsupported Attachment Type, SIZE=0	Informational
File is password protected	The file is password protected. or encrypted (0x000B).	ERROR
File is corrupt	The file is corrupt, so no content is indexed. Verify that the file(s) can be opened in their original application(s). Message detail may contain: File is corrupt (0x0009).	ERROR
File too large (no content indexed)	The maximum time allowed to extract content from a document was reached, mostly likely because it was too large.	ERROR
File too large (partially indexed)	The size of the document exceeded the product's maximum index-able file size, so the file was only partially indexed. This File Notice may contain one or more of the following details: • Following regions: { NEAContent,u_NEAContent } partially indexed	ERROR
Subdocument(s) not accessible	The file had references to other files that could not be opened/ accessed. This File Notice may contain the following details: • Supplemental files could not be opened (0x000C).	ERROR

File Notices

Error processing content	The application's 3rd-party content extraction application encountered a problem while indexing the document. This File Notice may contain one or more of the following details: • Exception occurred (0x03C0) • Access violation (0x03C1) • Integer divide by zero (0x03CD) • FileFilter is killing itself, or timing out • unknown error (0x0012)	ERROR
Unable to read attachment	This generally only occurs in Notes environments when The application encountered a problem accessing an attachment. It usually, but not always, means that the file was corrupt or empty. It could also be because the attachment filename is invalid. This File Notice may contain one or more of the following details: • Attachment read error • The system cannot find the path specified • The filename or extension is too long • Attachment name contains invalid characters • Attachment name property not set • Failed to open attachment • Error processing the attach-ment, an exception occurred	ERROR
Email container attachment	An attachment to an email was a PST or NSF container, so its content was not processed. (The file is a PST or NSF email source.)	ERROR

File Notices

Password-protected container attachment	An attachment to an email was a password-protected container file. This File Notice may contain the following details: • Container processing error - Reason: [#20116] Error extracting container attachment - [#20122] One of more files in the container is password protected.	ERROR
Attached or contained email processed as file	An MSG or EML was processed as a loose file document because it was an attachment. This File Notice may contain the following details: • MSG or EML is not processed as an email messages but content is indexed and searchable. • Msg or .eml message files found in containers or as attachments are processed as loose files instead of emails. Their content is indexed, but they cannot be searched by the usual email properties such as subject, sender groups, direction, etc.	Informational

File Notices

Container processing error	An error occurred extracting data from a container that was password-protected more than 10K documents in the container.	
	An attachment to an email was a password-protected container file. This File Notice may contain the following details: • Container processing error - Reason: [#20116] Error extracting container attachment - [#20122] One of more files in the container is password protected.	
	This error occurs when the container file type is not supported. The individual files within the container are indexed however. This File Notice may contain the following details: • Container file type not supported for extraction. The content of individual files within the container has been indexed but cannot be searched by file metadata properties	
	Error processing files within a container; process code returned -2.	
Language processing error	Language processing error (Language boundary identification failed.)	ERROR
MSG/EML File processing error as container	Unable to process MSG/EML which is an attachment to an email. Message detail may contain: • MSG/EML file processing error as container	ERROR
MSG/EML attachment processing error	Unable to process message attachment due to corruption, retriever is unavailable. These will be processed as loose files.	ERROR

File Notices

Embedded image in email	Attachment is an embedded image in email.	ERROR
File contains embedded content	File contains embedded documents.	Informational
File OCR'ed by the application	The application performed Optical Character Recognition on the file.	Informational
OCR conversion error	Error while attempting to use Optical Character Recognition to extract text from the file.	ERROR
File contains hidden content	File contains hidden content.	Informational
Hidden/embedded content check skipped	Hidden/embedded content check skipped as document pre-dates Office 97 or is otherwise unsupported.	Informational
File contains unknown embedded content	Embedded objects of unknown type were not extracted.	Informational
Attachment name not available in the message properties. Defaulting to the attachment display name	The name of the attachment could not be found in the message's properties. The default is to the display name of the attachment.	Informational
File has preset print area	Occurs when Microsoft Office documents, such as Excel spreadsheets, which contain a pre-defined print area selection cannot be fully represented or rendered in the output document.	Informational
IRM-protected document decrypted by the application	Occurs when application encounters an IRM-protected document, then proceeds to decrypt the document.	Informational
IRM-protected document decryption failed	Occurs when application encounters an IRM-protected document, but fails to successfully decrypt the document.	ERROR
Other warning	Appears when all other warnings do not accurately classify the specific error based on the available criteria.	ERROR

Case Administration

For information about how to maintain cases, refer to the following topics:

Selecting a Case

Case Admins who have been assigned to multiple cases but do not have the **System Manager** role are shown a list of their accessible cases when they first log in. Selecting a case displays the View Case Status page (see *"Monitoring Source Processing Status" on page 107*). For **Group Administrators**, the case list is not shown. **Group Admins** can view all cases, or create a new case within their group.

Note: If you are using the native viewer in eDiscovery Platform 9.1 or 9.5, you must first upgrade to eDiscovery Platform 10.0 and then use the Imaging Tool Upgrade feature for the existing cases.
You must upgrade a case using the Imaging Tool Upgrade support feature to be able to perform imaging-related operations in that case. See the *Imaging Tool Upgrade Guide* for details.

To select a current case

1. On the top navigation bar, click the drop-down and select a case. After you select a case, clicking **Case Home** displays the overall status for the case.

   Note: After selecting a case, if you see the message "*The digital fingerprint of emails processed into this case has to be updated because of the upgrade to Notes 10 or Office 2019. Please initiate the upgrade by navigating to "Update checksum for emails" within System / Support.*", refer to the *"Case Administration Workflow Recommendations" on page 18* section before running the update email checksums job.

2. To search the currently selected case, click the **Analysis and Review** module. To search a different case, select a different case from drop-down menu in the navigation bar.

Changing the Case Settings

Only a user with the **System Manager** role can change settings to cases after creation, including the case name, description, type, business unit, team members, and dates. To change the document sources for a case, see *"Selecting Document Sources and Pre-Processing" on page 49*. Case settings (depending on options available to modify) can be changed before or after processing on the **Settings** screen, either from **Case Home** (before processing), or from the **Processing** module (after the case is processed).

Note: You must have the **System Manager** role to change settings for a case.

To change settings for a case

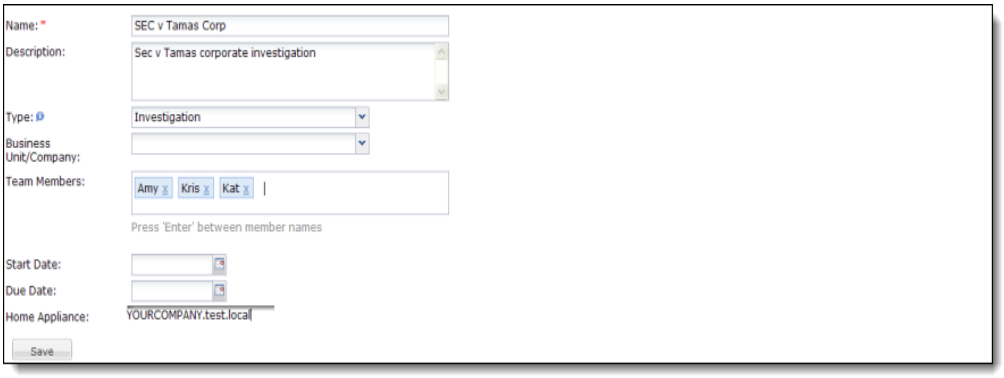
1. Navigate to the appropriate Settings screen, based on your case status:

   A. If the case has not yet been processed, on the top navigation bar, click **Case Home > Settings**. See the table *"Changing Case Settings Before Processing" on page 180*.

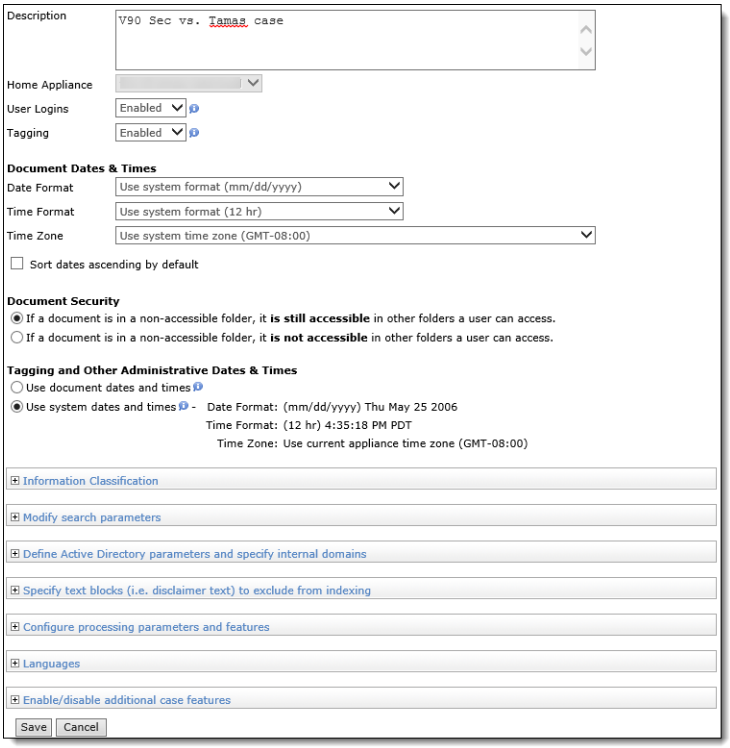
   B. If the case has already processed, on the top navigation bar, click **Processing > Settings**. See the table *"Changing Case Settings After Processing" on page 180***.**

   Note: To check the status of your case, click **Processing > Processing Status**.

   For cases not yet processed, the **Case Home > Settings** screen displays basic information for the case. Continue to step 2A.



For processed cases, the **Processing > Settings** screen displays detailed information for the case. Continue to step 2B.



2. To change case settings:

   A. Before the case is processed, change or complete the following available fields. An asterisk (*) indicates a required field. Fields that cannot be changed are indicated.

Changing Case Settings Before Processing

Field	Description
Name*	Change the case name (up to 35 characters).
Description	Change or enter a description for the case (up to 255 characters).
Type*	Change the type of case from the drop-down menu. **Note:** Users with the **System Manager** role or the **group administration** role can add or edit case types in the **All Cases > Settings** screen.
Business Unit/ Company	Change or enter the company's business unit or name to be associated with this case. **Note:** Users with the **System Manager** role or the **group administration** role can add or edit business units in the **All Cases > Settings** screen.
Team Members	Add or remove the names of users or team members involved with accessing and managing this case. The application automatically adds new names to the list, and can be re-used for future cases. **Note:** Team members can either be pre-defined by the **System Manager** role or the **group administration** role, or entered to create new members in the **All Cases > Settings** screen.
Start Date and Due Date	Change or enter the dates indicating the official start of this case, and the target due date for completion.

   B. After your case has already processed, change or complete the following available fields. An asterisk (*) indicates a required field. Fields that cannot be changed are indicated.

      › Click each category of case settings to view or change the current values. The following table describes each group of settings.

Changing Case Settings After Processing

Field	Description
Description	Enter or change a description of the case (up to 255 characters), even if you already entered one on the previous screen.
Home Appliance	(Appliance cannot be changed.)
User Logins	Select **Disabled** to prevent non-administrative users from accessing the case. You can enable user access after the initial configuration and indexing are complete.
Tagging	Select **Disabled** to prevent all users from tagging documents in the case.

Changing Case Settings After Processing

Field	Description
Document Dates and Times	Document-specific date/time settings are useful when the documents in a case originate in a different time zone from the location of the appliance. Each case can have its own document date and time settings, thereby allowing a single appliance to support cases originating from multiple locations. For example, a law firm headquartered in New York, which has its system date and time settings set to a US date format and Eastern time, may be managing a case with documents that originated in London. The system time zone is U.S. Eastern time and the format is based on the 12-hour clock. To enable reviewers to see document dates and times as the London custodian would see them, the administrator configures the following document settings: • **Date format**—dd/mm/yyyy • **Time Format**—24 hour • **Time Zone**—GMT With these settings, all document-specific information in the case is displayed in the document (London-GMT) time zone using the 24-hour clock. In addition, the European date format (dd/mm/yyyy) is used for displaying and printing reports. Select **Sort dates ascending by default** if you want all documents to be sorted in ascending date order and set as the default.
Document Security	Change security permissions for viewing documents in this case: • **If a document is in a non-accessible folder, it *is still accessible* in other folders a user can access**—(Default) Least restrictive: Allows users to view a document if the document is in a folder that they have permission to view (regardless of whether the same document exists in another folder that users do not have permission to view). • **If a document is in a non-accessible folder, it *is not accessible* in other folders a user can access**—Most restrictive: Prevents users from viewing a document if the document is in a folder that users do not have permission to view (regardless of whether the same document exists in another folder that users do have permission to view).

Changing Case Settings After Processing

Field	Description
Tagging and Other Administrative Dates and Times	Change whether dates and times are the same for case administration functions as for document display. Choose one of the following options: • **Use document dates and times**—Ensures that *all* date and time settings for the case (for administration and document display) are in the document format and time zone, as specified in the previous entry in this table. • **Use system dates and times**—Uses the system date and time settings for case administration tasks (such as user login tracking and export). Refer to *"Managing Schedules" in the System Administration Guide* for information on the system level date and time settings. Using the New York/London example (from the Document Dates and Times description), the administrator would choose **Use system dates and times** to keep administrative operations in the New York time zone (the system level time zone). However, if the all of the case administration and document handling were performed in London, the administrator would choose **Use document dates and times**.
Information Classification	
Enable automatic classification of incoming data. Note: Only policies enabled in the Information Classification portal will be utilized for classification.	Check to enable Information Classification in the platform. By default, Information Classification is disabled. **Note:** You must have enabled policies on the Information Classification portal side before enabling Information Classification in the eDiscovery platform. For more information, see *"Information Classification" on page 81*.
Modify search parameters	
Minimum size of document to return...	Specify the smallest size that a document must have to be returned in a search (10 KB default). Documents containing no indexed text may contain significant content, such as images. When searching for documents with no indexed text, you may want to ignore documents whose size is very small, which are likely to contain only insignificant images like those of a signature.
Maximum result size (documents)	Enter the maximum number of documents (100 to 10,000,000) that can be retrieved by a search (default is 1,000,000).

Changing Case Settings After Processing

Field	Description
Find Similar Settings	Set the default document similarity threshold. This is the setting used in the similarity histogram as the default "Minimum Rating" value. A lower value indicates items which are less similar (versus a higher value indicating closer similarity, nearly duplicate) to the original item. **Note:** During review, users can adjust this similarity threshold for any original item to find similar items for analysis. For more information, refer to *"Viewing Search Results" in the Veritas eDiscovery Platform User's Guide*. Adjusting the similarity settings does not require post-processing to be re-run. You can also set where similar items are found: across the entire case or within search results.
Define Active Directory parameters and specify internal domains **Note:** You cannot modify these settings after the case is created.	
Use Global Participants and Domains	Indicates whether the participant list includes participants obtained from an Active Directory server or is limited to participants discovered in the document sources assigned to the case (cannot be changed). IMPORTANT: There may be distinct differences as to how participants and domains are resolved depending on whether this setting is checked or not. This setting may also affect participants, filter counts and search criteria. For more information, see *"AD Synchronization and "Use Global Participant and Domain" Case Parameter" on page 45*.
Internal Domains	To add a domain specific to this case, enter the domain name and click **Add**. To change a domain name, select the domain, enter the correct name, and click **Replace**. To delete a domain, click the trash icon for the name.
Specify text blocks to exclude from indexing	
Indexing exclusions	To exclude commonly found blocks of text from the index, enter the text on one or more lines, and click **Add**. To change a text block, select the text block, enter the correct text, and click **Replace**. To delete a text block, click trash icon for the text block. The specified text is excluded from documents processed in the future, but is not removed from the current index.
Configure processing parameters and features	
Estimated number of documents in index	Enter the estimated number of documents to be indexed (100,000 to 10,000,000). Used only to optimize performance (not a hard limit).
Messages with no Sender email address	Select one of the following: • Process and set sender to "none." Process the message and assign the value "none" to the **Sender** field. • Process and set sender to last modifier. Process the message and assign the email address of the last person who modified the email in the **Sender** field. • Do not process. Do not include the email in processing.

Changing Case Settings After Processing

Field	Description
Enable Predictive Coding	Select the check box to enable predictive coding, the ability to learn the review criteria of your case and assess the corpus for relevant documents. For more information, see the *Transparent Predictive Coding User Guide*. **Note:** To enable predictive coding, you must also select the **Enable review, redaction, and production features** option under the Enable/disable additional case features section.
Extract email files to (Default directory is given)	Specify the parent directory to which you want to extract PST and NSF files when found inside container files (such as .ZIP files). This parent directory will contain a case specific folder (named for the case ID) when the case is created; this folder will ultimately contain the extracted files.
Extract documents from container files	Select the check box to have the system extract all files from the container or archive files, such as .zip files found as attachments in messages from PST, NSF, EMLX/EML/MSG sources. After files are extracted, the original container/archive file is excluded from the search results.
Convert supported mailbox files to PST	Specify the location (directory) where you want the system to convert other mail format files to PST and store them. **Note:** This folder is not automatically backed up with case backup.
Crawler Properties for Non-Email Items	Email messages are always indexed for all document sources. **Note:** These properties will be locked once Processing begins.
Process loose files that are 0 bytes long	Select the check box to process files that are specified as 0 bytes in size.
Process truncated Lotus Notes documents	Select the check box to process Lotus Notes files that are truncated due to excessive length.
Document duplication in milliseconds	Selected by default, this option allows de-duplicate documents based on the sent date of the document, in milliseconds (rounded up to the nearest second). (This option cannot be changed after processing. If the check box was not selected, duplicate documents were processed; though only the seconds value was used; not milliseconds). **Note:** This applies to both loose files and e-mail, and can only be configured or modified prior to processing.
Interpret ambiguous "##/##/##"-formatted dates for derived emails as if formatted as	Select the date format for ambiguous dates (mm/dd/yyyy versus dd/mm/yyyy). A derived email is the text content of an email that is enclosed within another email. The application uses these emails to construct more complete and accurate discussion threads. However, because derived emails are text only, there can be ambiguities in how the application should interpret the sent date of the email.

Changing Case Settings After Processing

Field	Description
Process a ".TIF" file's matching ".txt" file	A TIF/TXT pairing is produced when documents are in imaged form (for example, scanned from paper documents). If optical character recognition (OCR) is applied to extract the text, the result is a pair of files that represents the content: an image (TIF format) and its extracted text (TXT format). The following options are supported. • Never. Process all ".TIF" files as regular image files, independent of matching ".txt" files. Do not perform any special actions when processing the file. • When the ".TIF" file is found in the specified folder and the matching ".txt" file is found in the specified folder. The system searches for a .txt text file that has the same name as the TIF file (such as "memo.tif" and "memo.txt") and is in the same folder. If the text file is found, it is processed instead of the TIF file. • When a pair is found within the same folder. The system searches for a .txt text file in the specified folder that has the same name as the TIF file in the other specified folder. If the text file is found, it is processed instead of the TIF file. • As described by a mapping file at the root of the source. The system searches for a text file that is mapped to a TIF file with the name that is found in the root folder of the source. If this mapping file is found and the corresponding text file is found, the text file is processed instead of the TIF file.
Specify a filter to use when excluding known files	Select the list to exclude known files during indexing. See *"Setting Pre-Processing Options" on page 74*. This list will be used to exclude known files during indexing. In addition to the default Veritas NIST list, custom lists can be defined by the System Manager in the "Known File Lists" area under System > Known Files. **Note:** The selected list cannot be changed after indexing has begun
Hidden, Inserted, and Embedded Content	By default, the application finds and indexes all text contained within a document. However, if the text is obscured or hidden, it can be difficult to find and view the indexed text. Identifying content enables you to search and filter for hidden and embedded content. Extracting embedded content enables you to view embedded documents as attachments or embedded content. • Identify and extract different file types. • Identify only. • Don't identify or extract. Text is indexed, however, content might not be viewable if the information is not identified or extracted. **Note:** These properties will be locked once Processing begins.
OCR Processing	

Changing Case Settings After Processing

Field	Description
Use Optical Character Recognition (OCR) for documents where no text is found	Choose whether to process image and non-text files without OCR. If you enable OCR, select the file types to process when no text is found. By default, OCR is disabled. **Note:** Processing case files requires more time when OCR is enabled. It is strongly recommended that you leave this option disabled, with the exception of only very small cases. For normal-sized cases, leave this option off. Later, you can perform a search to select the documents you want to process with OCR. For more information, see *"Processing (or Resubmitting) Documents for OCR" on page 114*. **Note:** As of version 7.0, the application can OCR process documents that are in the Icelandic language.
Languages **Note:** You can change all language settings after initial processing and then rerun post-processing.	
Automatically identify the following languages within your case	Select check boxes to specify the languages that you want to include in document searches. Select only the languages that you believe may exist in your case. Languages that are not selected will not be automatically identified and will be classified based on the settings below. The most commonly-spoken languages are selected by default.
When a portion of a document can be interpreted as more than one language	Sometimes the same words and characters are used in more than one language. This setting helps to accurately identify these shared words or characters. Specify the precedence order for determining the language (Chinese, Japanese, and Korean only). Click the **Move Up** or **Move Down** buttons to change the order. You cannot modify these settings after the case is created.
For documents that can not be automatically identified	Select the single language to apply from the drop-down list if it is not possible to identify languages in a document automatically. For example, it is difficult to accurately identify documents with limited content, such as short emails and appointments. If the expectation is that your data set is mostly in one language, such as English, then configure this setting to that language to best classify these documents. Alternatively, you can classify these documents as "Other."

Changing Case Settings After Processing

Field	Description
Advanced Options	For small amounts of document content, it is not possible or desirable to automatically identify the language. You can configure the minimum number of characters and the percentage of a document's content that is required to automatically identify a language within the document. Exceeding either the character or percentage threshold will trigger automatic language identification. When you click the **Advanced Options** button, the Automatic Language Identification Advanced Options window opens. Configure the following settings: • Specify the minimum number of characters to automatically identify a language (default is 200). • Specify the minimum percentage of a document's content to automatically identify a language (default is 10%). • For content that does not meet the other thresholds or cannot be automatically identified for any other reason, choose a language for manual identification.
Enable stemmed search for the following languages	Select check boxes to enable stemmed searches for specific languages. A stemmed search automatically finds documents that contain common variations of a word that is entered as part of a query. For example, if you search for the word "test," a stemmed search also finds variations such as "testing,""tests," and "tested." Two English options are available to support stemmed searches. Both are selected by default: • English—Uses a sophisticated linguistic stemming algorithm to determine stemming rules. For example, this option considers "went" as a variant of "go." • English (suffix-based stemming)—Uses the Porter algorithm to strip out common word suffixes (such as "s" or "ing") for stemming. This algorithm is useful for finding nouns in their plural and singular forms. **Note:** Each additional language increases processing time within your case.
Monthly Billing Model	
Standard or LIHO (Low-In/High-Out)	The option selected at case setup is shown, however these settings cannot be changed. See *"Defining New Cases" on page 18* in the table: "New Case: Processing Settings". **Note:** The LIHO billing option only appears if you have a consumption based license.

Changing Case Settings After Processing

Field	Description
Enable/disable additional case features	
Enable advanced processing options configuration (also known as pre-processing)	Select the check box to allow use of the pre-processing features. See *"Pre-Process Your Source Data" on page 67*.
Enable review, redaction, and production features	Select the check box to allow use of the redaction feature. See *"Setting Up Redaction Sets" on page 138*. **Note:** This feature must be enabled for Predictive Coding to function.
Enable email header viewer	Enables the viewing of email headers in email messages. By default, this option is disabled.

3. Click **Save** to submit the changed profile, or click **Cancel** to discard your changes.

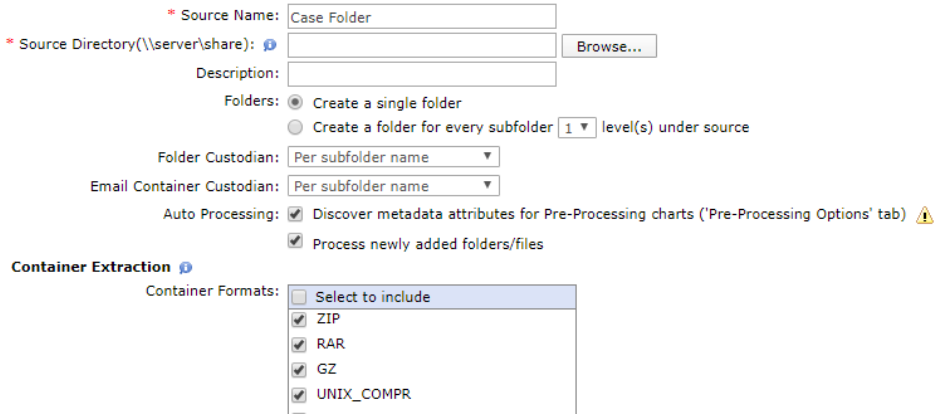
Changing Custodian Assignments (on Newly Discovered Data)

If you have new data that was discovered in a case but need to change the default custodian, you can use the Manage Sources page to select a case and change custodian assignments.

Note: Changing the custodian in the Edit Case Folder Source page only applies to newly discovered data in the case folder. (Be sure to run discovery first before changing the custodian assignment for that data.)

To change the custodian assignment in a case folder (for newly discovered data)

1. On the top navigation bar, for a selected case, click **Processing > Sources and Pre-Processing**.
2. Click to select the source you want to edit.
3. On the Edit Case Folder Source screen, change the Folder Custodian by selecting another name from the drop-down list.



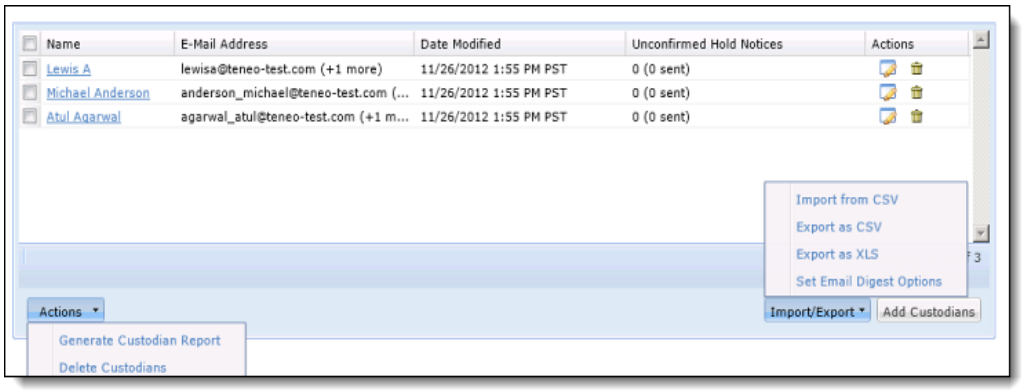
Note: Your change will be saved, however the new custodian will not be reflected in the Processing Options chart unless you change the custodian from the main Sources page.

4. Click **Save** to change the custodian, or click **Cancel** to discard your changes.

Note: For your changes to take effect, you must run post-processing.

Managing Custodians

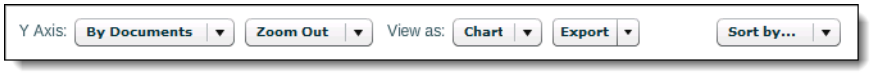
A centralized Custodians menu allows you to add, import custodians, and destinations to your data map, and customize, analyze and configure employee identification information. You perform these tasks from the **Case Home > Custodians** menu. This menu provides a unified view and workflow for handling custodians and custodian data. For more information on Custodian Management, see *"Collection Workflow" in the Identification and Collection Guide*.



Analyzing Case Data

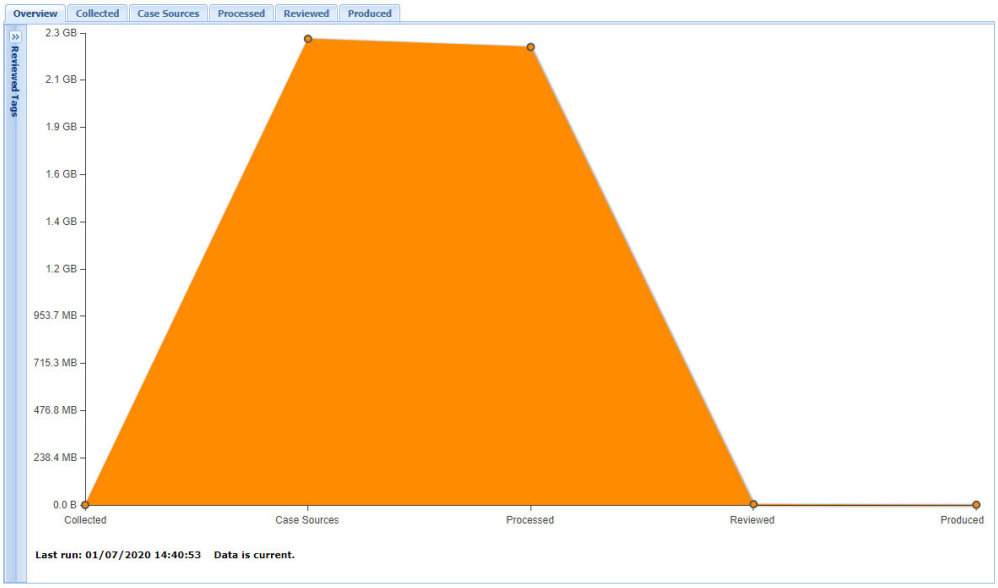
Viewing analytics across the lifecycle of your case is simple. From Collections through Pre-Processing, Processing, Review and Production, you can see data that spans all stages of the case. These charts are available from **Case Home > Data Analytics**.

The criteria for each tab (except the Overview tab) can be set through a group of drop-down menus along the top of the chart. Charts are redrawn to meet your criteria. You can also export the charts in table format as either CSV or XLS files.



There are six tabs available in Analytics:

Overview tab



The Overview tab gives a view of all data across the five stages of a case. Hovering over any one of the five data points will give you details about that stage. Clicking on a data point will give you a drill-down view for the data in that stage.

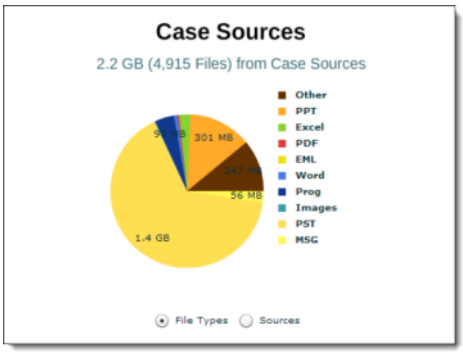
Note: In the Overview tab, the "Reviewed" stage is can be customized depending on which stage of review you want to track. Click on the Reviewed tags panel (which is collapsible button) to select which tags you want to display in the Reviewed column of the Analytics chart.

Collected tab

Note: This tab is only available if you have purchased the Collection module.

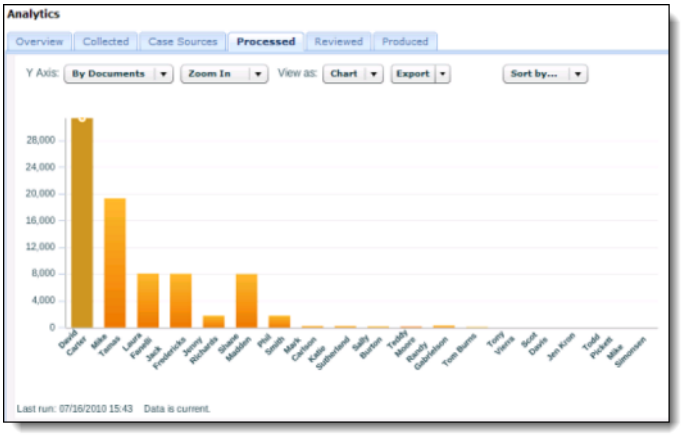
The Collected tab shows the data volume (GB volume or File count) of data collected with the Collection module. Data is presented by Custodian, and clicking each Custodian column will bring up a drill-down view of the File Types and Sources of data collected for that Custodian.

Case Sources tab



The Case Sources tab displays the volume of data that has been added to a case for discovery. This includes Collection Sets, as well as other Case Sources (Case Folders, EDRM Sources, etc.). Individual drill-down charts per custodians are also available. This shows the data volume and item count after extracting container files (such as ZIP files and PST files).

Processed tab



The Processed tab displays the volume of data actually Processed for this case, organized by Custodian.

Reviewed tab

The Reviewed tab reports the volume (or document number) of items that have been tagged by reviewers. You can custom select which tags to report on via the Reviewed tag icon in the Overview tab. As with all of the other reports, individual drill-down views are available per Custodian.

Produced tab

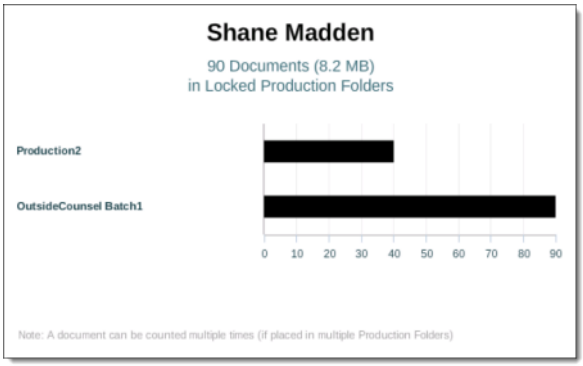


Chart of an individual custodian's produced documents

The Produced tab reports on the volume of items that have been Produced in Locked Production Folders. The individual drill-down for each Custodian will show the volume in each Production Folder.

Configuring Review Dashboard Statistics

The Review Dashboard is populated with case report data. By default, this data is collected automatically and reports are generated every 24 hours.

To set up dashboard statistic or case reports

From **Case Home > Case Reports**:

- Enable or disable scheduled jobs that collect and analyze case data by selecting or deselecting the **Enable Scheduled Jobs** option.

  Every four hours case data is checked to ensure the case reports are current. If the reports have become stale, a new job is immediately scheduled and run. The job retries until it is able to complete successfully.

  These jobs are run in the background and are not visible from the **System > Jobs** or **Schedules** pages. any jobs that fail are displayed in the Jobs window.
- Start a collection job manually by clicking the **Start Collecting** link.
- Set the tags used during report generation (up to 50) by expanding tag sets and selecting or deselecting specific tags.

For more information about using the Review Dashboard, refer to *"Accessing the Review Dashboard" in the User Guide*.

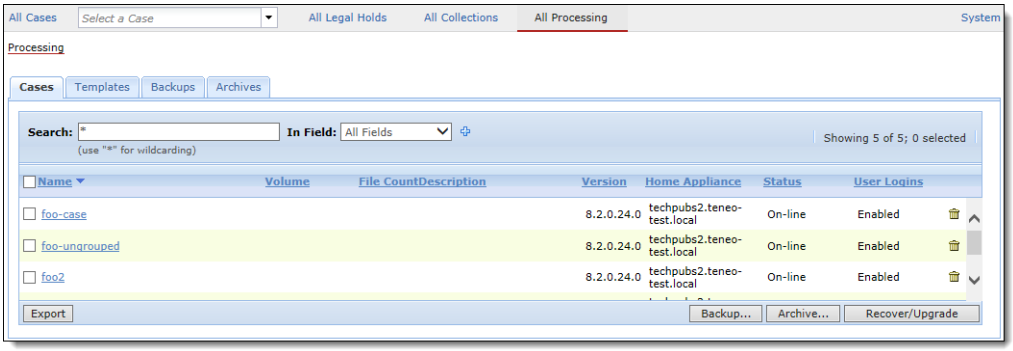
Managing Cases

On the Manage Cases page, users with the **System Manager** role can view the status of all cases, add new cases, change case configurations, backup and restore cases, and define templates to streamline the creation of new cases.

To view or change the current cases

1. On the top navigation bar, click **All Processing**.

   The Cases tab displays on the Processing screen showing status for all cases.



   The cases list indicates the number of documents indexed for each case, the number of bytes crawled thus far to create the index, the home appliance where the case resides, the case status, whether non-administrative users can access the case, and the available action for each case.

   Note the following case status values:

   - **On-line.** The case is active and available for user access. Click the status link to view the current case status.
   - **Off-line.** The appliance where the case resides is off-line (refer to *"Enabling, Disabling, and Restarting Appliances" in the -System Administration Guide*).
   - **Unavailable.** The case cannot be accessed. Try restoring the case from your latest backup (refer to *"About Restore" in the System Administration Guide*).
   - **Invalid Case.** The integrity of the case has been compromised and Additional Processing option is no longer available. Navigate to **All Processing > Backups** to restore the case.
   - **Recoverable.** A minor error has occurred that requires you to recover the case. Navigate to **All Processing** to recover the case.
   - **Backing up.** The case is currently being backed up.
   - **Restoring.** The case is currently being restored.

- **Recovering.** The case is in the process of being recovered.
- **Deleting.** The case is in the process of being deleted.
- **Processing.** The case is being indexed.

2. To add a new case, and specify the case name and document sources (see *"Defining New Cases" on page 18*). To view processing status for a case, click the case name (see *"Monitoring Source Processing Status" on page 107*).
3. To export the case list in CSV format, click **Export**, click **Save**, and specify a location.
4. To define templates for creating new cases, click the **Templates** tab (see *"Defining Case Templates" on page 198*).
5. To back up or archive cases, refer to *"Creating Case Backups" in the System Administration Guide*.
6. To recover cases after a system failure, or upgrade a case, click **Recover/Upgrade**. A successful recover operation restores the selected cases automatically, including tag information. Information is recovered from current on-disk state of the case, not from a backup.

   Note: You can restore backups and archives that are selected from the **Backups** or **Archives** tab, in which case the case reverts to the content of the backup or archive. If you recover a status with the state "recoverable," the data in the case is not modified. The system makes only the changes that are required when a Secondary appliance joins a cluster or a primary appliance is rebuilt from backups.

7. To delete a case, click trash  icon for the case.

Viewing Case Status Report

Starting with release 10.0, users with the **`Allow Support Access`** permission can generate a case status report using the "Case Status Report" support feature. This report helps administrators to see active and inactive cases, and decide on when to archive an inactive case.

A user with the **`Allow Support Access`** permission can generate a xlsx report for all available cases in the system with following columns:

- **Case** – The name of the case.
- **Created By** – The name of the user who created the case. If the full name of the user is not available, then the username is displayed.
- **Creation Date** – The date (MM-DD-YYYY) and time (HH:MM:SS) when the case was created. The time zone is also displayed.
- **Days Open** - The number of days from the case creation date to the current date in days.
- **Last Activity Date** -The date on which the last processing or search was performed on the case.

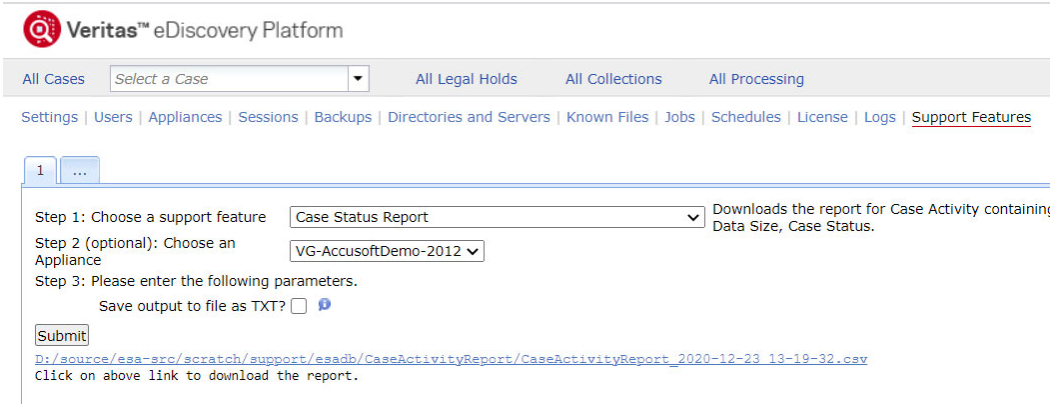
- **Days Inactive** - The number of days from the "Last Active Date" to the current date in days.
- **Processed Data Size** – The size of the processed data for the case.
- **Status** – The status of the case, it can be **Active** or **Inactive**.

When a case has no processing or search activity for 30 days, it is marked as an inactive case by default. This duration can be configured using the **`esa.support.caseactivityreport.threshold.days`** property. The default value for this property is set as 30 days.

The case status is Active when the Days Inactive is less than or equal to the days defined for the above mentioned property, and Inactive when the Days Inactive is greater than the days defined for the above mentioned property.

To generate a case status report

1. Using an account with System Management permissions, log onto the eDiscovery Platform web interface.
2. From the **System** > **Support Features**, select **Case Status Report**.



3. Optionally, choose an appliance to which the feature will apply.
4. Click **Submit**. A download link appears on successful execution of the feature.
5. Click the link to save the report on your computer.

Defining Case Templates

If you routinely create cases with the same advanced settings, folders, tag categories, and/or groups, users with the **System Manager** role or the **Group Admin** role can define the most common settings in a case template, and then use the template to create new cases. To use a template to create a new case, see *"Defining New Cases" on page 18*. Note that templates are stored on the primary appliance.

To define case templates

1. On the top navigation bar, click the **All Processing** module.
2. Click the **Templates** tab.
3. To add a new template and specify the advanced case settings.
   A. Click **Add** and specify the following information. An asterisk (*) indicates a required field.

New Template Name

Field	Description
Name*	Enter a template name (up to 35 characters).
Description	Enter a description of the template (up to 255 characters).

   B. Click each category of case settings to view or change the default values.
   To view all the possible settings, refer to the table on *"Changing Case Settings After Processing" on page 180*.
   C. Click **Save** to submit the new template, or click **Cancel** to discard your changes.
4. (Optional) Add folders or tags within the case template by navigating to the Folders or Tags page respectively after clicking **Save**.
5. To delete a template, click trash icon for the name on the template list.

Producing Search Results

Production is the stage in the eDiscovery process in which documents are prepared for delivery to a requesting party. The production feature allows you to prepare documents for delivery.

Document production occurs in the following stages:

1. **Folder creation**

   To produce a set of documents, the documents must be copied into a production folder. To set up production folders, see *"Setting Up Production Folders" on page 129*.

2. **Document selection**

   Select the documents that you want to produce and copy or move them to a production folder. You can move or copy documents in either of the following ways:

   - Use the Action menu in Review mode, as described in *"Using the Document Review Screen" in the User Guide*.
   - Use the tagging window, as described in *"Bulk Tagging Reviewable Items" in the Veritas eDiscovery PlatformVeritas eDiscovery Platform User's GuideUser Guide*.

3. **Production**

   To produce the set of documents, see *"Running a Production" on page 199*.

4. **Production Review**

   Review the production in Review mode, as described in *"Using the Document Review Screen" in the User Guide*.

5. **Production Export**

   Export the produced documents from Review mode, as described in *"Performing Exports" in the Export and Production Guide*.

When generating the production document, the application will automatically scale down the original page to allow for the specified header and footer to be displayed without encroaching on the original document content. This is the case for all document types, including TIFFs. For example, if you have processed an already-numbered TIFF into the application, on production you will see the application-generated production number at the top of the page and the original TIFF production number in the TIFF region of the page.

Running a Production

The application provides three options for production: image, near-native, and mixed. Mixed-mode productions enable you to choose the file types that you want produced as images and the file types that should be produced in their native format.

Mixed-mode productions become useful when you want the majority of your files to be produced as TIFFs, but prefer to produce Excel spreadsheets natively, since they are fairly complex to render as TIFFs.

The application allows you to run multiple production jobs simultaneously. These production jobs run until they complete enabling you to start a production job and run it overnight without needing to monitor the jobs status. If the job fails, it is restarted.

Note: If the case is created in pre-10 release, then after upgrade to release 10.0, production folders of type Images and Mixed are affected in Imaging Tool Upgrade. For Native type production folders, all operations are available before, during and after the Imaging Tool Upgrade is performed. See the *Imaging Tool Upgrade Guide* for details.

To run a production

1. On the top navigation bar, for a selected case, click **Analysis and Review > Folders**.
2. Select the production in the **Production** folder and click the "Edit" icon, and select **Edit**.
3. In the Production Folder window, verify the production folder settings are set correctly.

   For information about production settings, see *"Setting Up Production Folders" on page 129*.
4. When ready, from the main screen (or from Production Folder window), click **Lock/ Produce.**

   This produces the documents and locks the folder so that no more documents can be added to the folder. To reverse this operation and make the documents unavailable for production, click **Unlock/Unproduce**.
5. After the production is run, the document is available for viewing in Review mode prior to export. The user opens Review mode, and chooses **Productions** from the **View** menu.

To use slip sheets for productions (and view details of slip-sheeted documents)

1. From the **Analysis and Review** module, click the drop-down showing "All Documents" (default).

   Or

   From the **Analysis and Review** module, select the **Folders** tab.
2. Select the Productions folder, then click **Edit** from the Actions menu. The **Production Folder** menu displays.
3. On the bottom line of the menu, click the slip sheet report icon  to launch creation of the report (zip file).
4. From the jobs window, select the report zip file to download and review the errors associated with the production.
5. To export the production, select **Export > Production** and specifies export parameters.

Note: In versions 5.5 and later, you can unlock/unproduce a production before or after export, as long as it is not currently running, or not in a failed state. If the production is running, it must be stopped and the job deleted before the folder can be unlocked.

Unlock and Unproduce a Production Folder

Production folders can be unlocked and unproduced after the produced folder has been exported. This enables you to remove privileged documents that were inadvertently included in the original production, and then produce a new set of documents using the original production numbers.

You can unlock and unproduce a production when:

- You have permission to unproduce an exported production.
- The production is not running, pending, or in a failed state.
- The production job run has been deleted.

To unlock and unproduce an exported production

1. On the top navigation bar, for a selected case, click **Analysis and Review > Folders**.
2. Confirm that you want to unproduce the folder (and you have permission to do so) to remove all production information.
3. Select the Productions folder you want to unlock/unproduce, then click the "Edit" icon and select **Unlock/Unproduce**. (Alternatively, select Edit, then from the Production Folder window, click **Unlock/Unproduce**.)
4. If this production has been exported and is being stored on disk, delete the production export manually.

   Deleting the export prevents two documents from being branded with the same production numbers.

Reviewing a Production

After you create a production, review the documents to verify everything displays accurately.

During the production review, verify the following:

- Headers and footers display the correct information and are positioned appropriately.
- Redactions display properly.
- Production numbering is accurate.

To view production statistics

1. From the **Analysis and Review** module, click the **Folders** drop-down menu.

   This menu is next to the search field and displays "All Documents" by default.
2. Click to select the **Productions** folder, then click the **Edit** icon and select **Edit** from the Actions menu.
3. From the Production window, click the **Production** tab.

   The current production statistics display.
4. If the production is still running, click **Refresh** to refresh production statistics.

   Refreshing the statistics enables you to get an early view of any potential problems in the production by viewing the statistics for the number of slip sheets produced, for example.

To review a production

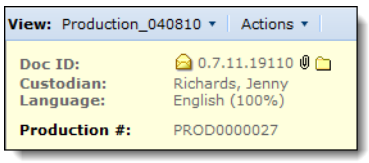
1. From the **Analysis and Review** module, select the production folder to review from the Folder menu.
2. Click **Go**.

   The produced documents display in List view with their production number displayed.

   Note: You can sort documents based on their production numbers.

3. Click the **Review Mode** icon to go to Review Mode.
4. From the View menu, click **Production** and select the production folder you want to review.

   The documents display.
5. Verify the following:
   - Headers and footers display the correct information and are positioned appropriately.
   - Redactions display properly.
   - Production numbering is accurate. By default, production numbers for documents and their attachments display in the header.



Managing Case Schedules and Jobs

The topics in this section describe how to manage case schedules and tasks, or view job details:

- *"Managing Case Schedules" in the next section*
- *"Managing Case Jobs" on page 205*
- *"Viewing Documents Processed for OCR" on page 206*

Managing Case Schedules

For each case, you can schedule tasks to automatically process the documents for one or more sources in the case. To add document sources to a case, see *"Managing Case Sources and Custodians" on page 49*.

To manage case schedules

1. On the top navigation bar, for a selected case, click **Case Home > Schedules**.
2. To enable or disable a task, select the check box next to the task(s), and click **Enable** or **Disable**. Click the first check box to select all the tasks.

   Note: The **Scheduled Time** column shows each task's first scheduled run time.
3. To schedule a document processing task:

   A. Click **Add** and specify the following information. An asterisk (*) indicates a required field. (The Scope will always be for the case specified.)

Schedule Properties

Field	Description
Description	Enter a description of the task (up to 255 characters).
Initial Run Date*	Specify the date of the first execution of the task: • Click calendar  icon and select a month and day. or • Enter the date directly as: MM/DD/YYYY. **IMPORTANT**: If you run discovery tasks for email server/archive sources, do not run the discovery and document processing tasks at the same time (see *"Discovering Archive Sources" on page 45*).
Start Time*	Enter the start time in 24-hour format (HH:MM).
Max Duration*	Select the **unlimited** check box, or enter the maximum number of hours that the task can run (1 to 500).
Stop Date	Specify the date of the task should stop: • Click calendar  icon, and select a month and day. or Enter the date directly as: MM/DD/YYYY.
Stop Time	Enter the stop time in 24-hour format (HH:MM).
Frequency	Select a recurring schedule (**Daily**, or **Weekly**).

Schedule Properties (Continued)

Field	Description
Enabled	Select the check box to enable the next scheduled execution of the task.
Sources*	Select one or more of the document sources in the case. Hold down the **Ctrl** or **Shift** keys to select multiple sources.

B. Click **Save** to submit the new schedule, or click **Cancel** to discard your changes.

4. To change a schedule, select the task type (the task type cannot be changed).
5. To delete a task, click trash icon for the task.

Managing Case Jobs

On the Jobs screen, as well as the Jobs window, you can stop unscheduled case jobs that are still running, and delete jobs that are completed or stopped. Unscheduled jobs include tagging, exporting, and printing documents.

Note: Scheduled tagging operations using SmartTags are displayed on the Manage Schedules page (see *"Managing Case Schedules" on page 204*).

For export jobs, users normally download the exported documents from the Jobs window as a single ZIP file. If the exported documents exceed the specified maximum size (refer to *"Defining System Settings" in the -System Administration Guide*), you must access the files directly on the appliance. Click the information icon in the Jobs window to view the export job ID.

The information icon text also includes the full file location (as described in the infobubble), which takes the following form:

<esa_root>\data\<db_name>\dataStore_case_<dataStore_id>\<user_id>\jobRun_<run_id>

To manage unscheduled case jobs

1. On the top navigation bar, for a selected case, click **Case Home > Jobs**.
2. To limit the list of jobs displayed, select a user or update time from the **User** and **Jobs updated** drop-down menus.
3. To view the job status log for a job, click job icon in the status column.
4. To stop a running job, select the check box next to the job, and click **Stop Jobs**. (To stop a job from the Jobs window, click stop job icon.) When a job is stopped, users can see the status change in the Jobs window. To open or close the Jobs window, click **Jobs** at the top of the screen (above the navigation bar).
5. To delete a completed or stopped job, click trash icon, or select the check box next to the job, and click **Delete Jobs**. Note that all files associated with a deleted job are also deleted, such as PDF reports and PST zip files, and will no longer be accessible from the Jobs window.

Viewing Documents Processed for OCR

Using the OCR batch function, you can submit selected documents to be processed for OCR after your case has initially been processed. To process documents with OCR and view results of the job, see *"Processing (or Resubmitting) Documents for OCR" on page 114*.

Managing Review Using Automation Rules

Introduced in version 8.0, this feature allows a **Case Admin** or **Case Manager** to use automated rules to streamline the review process. Under **Analysis and Review** is the **Automation Rules** menu where you can create, schedule and run rules. Actions run by Automation Rules will generate their own report.

This feature supports the following types of actions and workflow:

- Finding items using saved searches or using existing tags, then copying, moving them into and out of existing folders, and then applying tags.
- Users can create multiple action sets as a part of an automation rule and also schedule multiple rules to run sequentially.

Refer to theses Automation Rules topics

- *"Creating an Automation Rule" on page 208*
- *"Running an Automation Rule" on page 210*
- *"Viewing and Editing an Existing Rule" on page 210*
- *"Disabling an Existing Rule" on page 210*
- *"Generating Rules Reports" on page 210*
- *"Best Practices and Tips" on page 210*

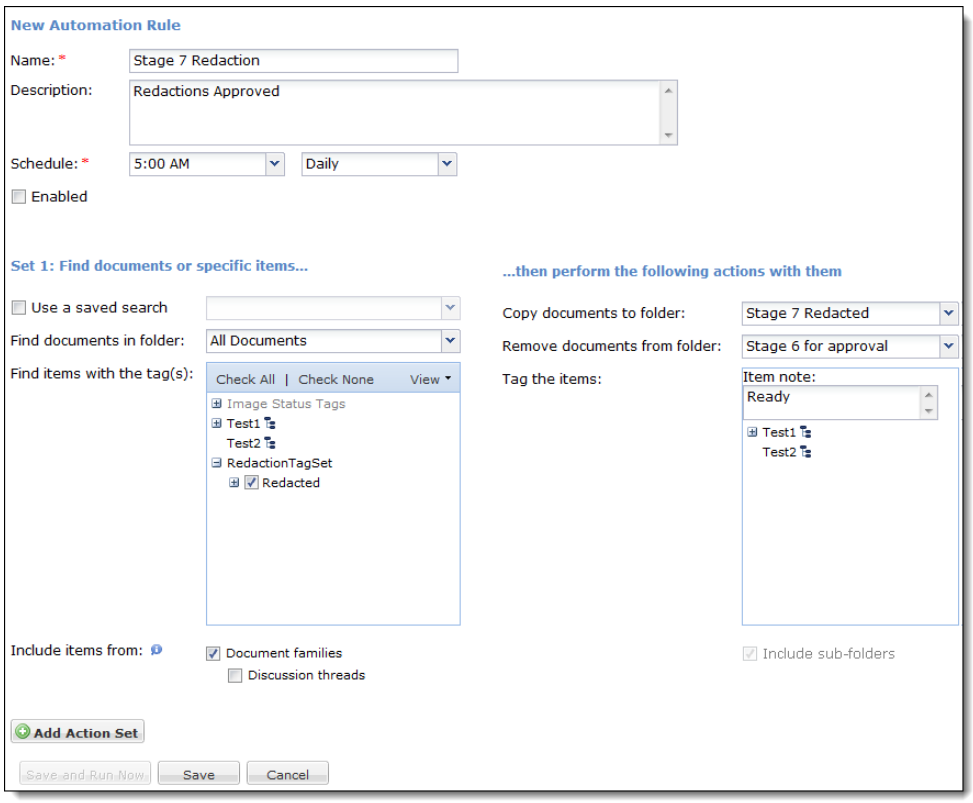
With the case selected, go to **Analysis and Review > Automation Rules** and click the **Create Automation Rule** button.

Creating an Automation Rule

You must have the **Case Admin** or **Case Manager** role, or have been granted the "allow access to automation rules" privilege.

A new rule has the following required fields:

- Name
- Scheduled time to run
- At least one action set in order to be saved



.

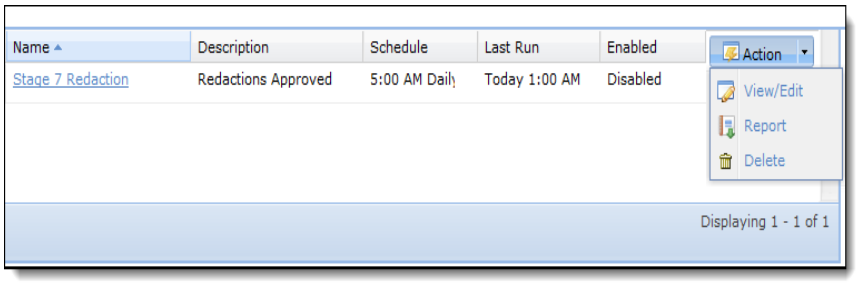
Automation Rules

New Automation Rule	
Field	**Description**
Name	Give the new rule a name.
Description	Enter a description for the automation rule.
Schedule	Select a time and frequency for the rule to run. You will not be able to save the rule without selecting schedule criteria. If you want to save the rule without running it, uncheck the **Enabled** check box, which is on by default.
Enabled	Save the rule as enabled (ready to run) or not. If you are adding more actions or are otherwise likely to edit, leave the box unchecked.
Find documents or specific items...	
Use a saved search	Check the box to use the drop-down list to locate a saved search.
Find documents in folder	Locate the items on which your rule is to run by searching for folders by name (for example: "Batch 12"), or by locating them from the selections displayed. "State" allows the **Case Admin** to see if the file is being reviewed, and to edit the folder name and description if not in use.
Find items with the tag(s)	You can check all available tags, or specific tags. The View drop-down list changes how the tags appear: **collapse all**, **expand selected**, or **expand all**.
Include items from document families	Include all document families and or discussion threads across the case that are related to the selected items. If the box is unchecked, this search will only apply to items. An **item** can be an individual email message, attachment, embedding. A **document family** is the parent document at all attachments.
Include items from discussion threads	A **discussion thread** is any email message and all replies and forwards.
...then perform the following actions with them:	
Copy documents to folder	Search by folder name or identify the folder from the selections. You can also create a new folder, or edit details for an existing folder.
Remove documents from folder	Search by folder name or identify the folder from the selections. You can also create a new folder, or edit details for an existing folder.
Tag the items	You can input a text item note, and chose any or all tags to apply from the selection give.
Include sub-folders	If there are no sub-folders, the option will not be active.
Add Action Set	A rule can consist of multiple action sets. They will run sequentially. **Note:** Up to 100 action sets are supported.
Save and Run Now Save Cancel	You must perform one of these actions to exit the screen.

Running an Automation Rule

When all the required fields are completed and the rule is validated, you will have options to "Save and Run Now" or "Save". Saving an enabled rule will run the rule at the time specified when it was created.

Viewing and Editing an Existing Rule



The **Case Admin** and **Case Manager** (or user with the "see automation rules" privilege granted) can see saved rules on the case's Automation Rules page. Scheduled time and frequency, last run date, and enabled status are all displayed. To edit or delete an existing rule, click the **Action** button on the far right, then select **View/Edit**, **Report**, or **Delete** from the drop-down list.

Disabling an Existing Rule

Click the rule's name or select the **View/Edit** action from the drop-down list. Uncheck the "Enabled" check box and save the rule.

Generating Rules Reports

On the Automation Rules page, click the **Action** button and select **Report** from the drop-down list. This will produce an Excel file containing the details of when the rule was created, its owner, user-specified description, schedule, whether it is enabled, and the number of action sets. The number of times the rule has been run is also reported.

Best Practices and Tips

- Schedule rules to match the rate of data influx or review workflow: they can be run once, nightly, hourly, or weekly.
- Do not schedule rules that operate on the same data set or tag set at the same time: the data set is locked while work is being performed.
- Users can see all the automation rules for the cases to which they have access listed on case's **Schedules** page. Users who have access limited to certain cases will not be able to see all rules for all cases. Task type, scheduled time, last run, description, and whether the rule is enabled are all shown on the **Schedules** page, but users other than **Case Manager** and **Case Admin** cannot edit, delete, or otherwise change automation rules.
- If another user deletes a folder or tag that is used by an Automation Rule, an error will display with the rule is run.

Using the Dashboard

The Dashboard provides a convenient, at-a-glance summary of case status. Reviewers and **Case Managers** can get live status of document and tag status and assignments. **Case Managers** can view the overall progress of a case and reviewer progress.

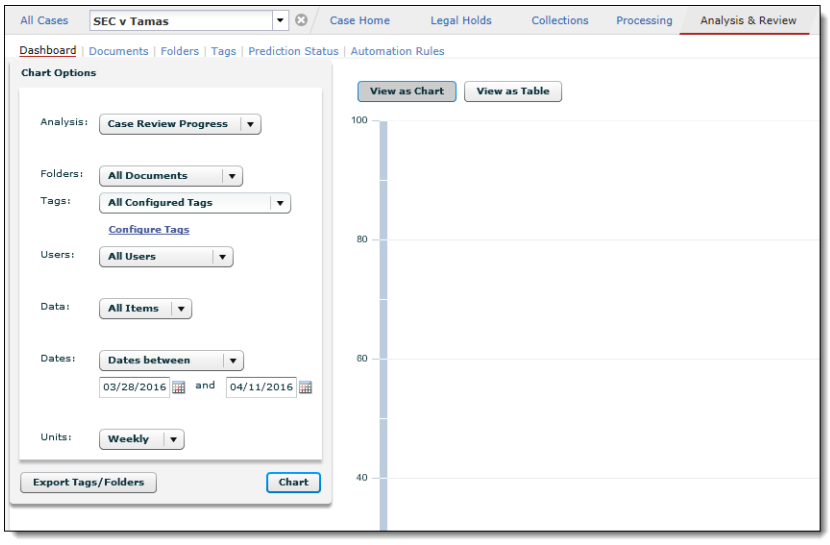
Refer to the following topics:

Access the Review Dashboard

All information displayed on the Review Dashboard is case-specific. Select the case from the drop-down on the navigation bar to view statistics for that case.

To access the Review Dashboard

1. If not already in the case, click the drop-down on the navigation bar to select the case you want to view.
2. Click **Analysis and Review** on the top navigation bar.



Selecting **Dashboard** opens the Chart Options menu on the left and the Chart (or Table) view on the right. By default, the information displays in chart form.

3. Select the report you want to view: click the **Chart** button to generate it.
4. To display the dashboard information in table form, click the **Table** button.
5. To search for the documents represented by a pie segment, a bar chart segment, or a number listed in the **Table** tab, click on the segment area or number.

   The search is performed, and the Search Results screen opens to show the documents that match. For information on the Search Results screen, refer to *"Viewing Search Results" in the Veritas eDiscovery Platform User's Guide*.

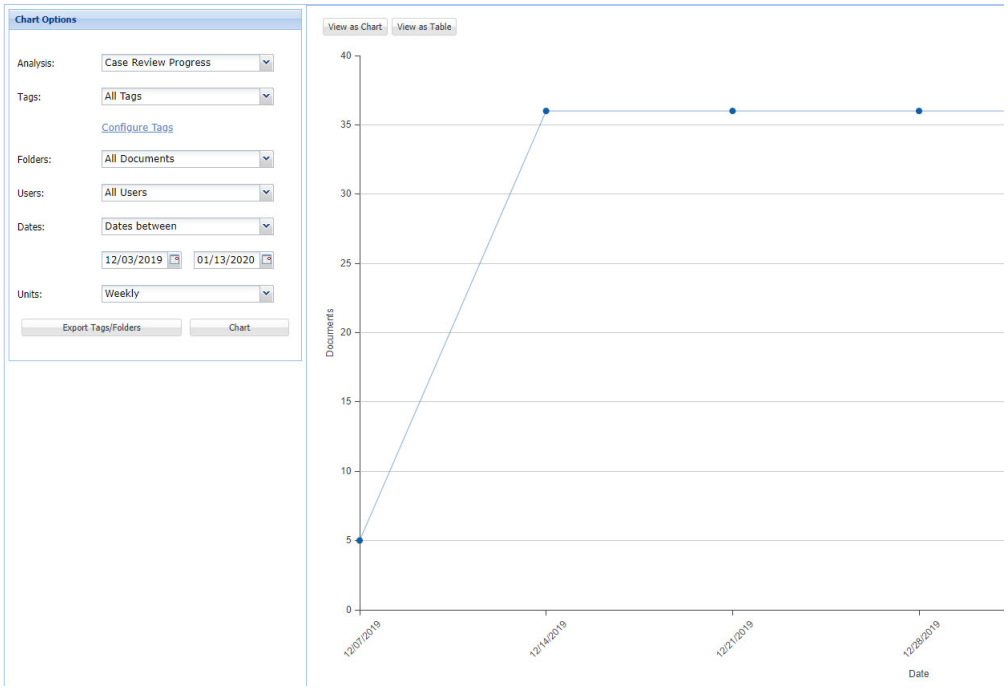
View Case Progress

The Case Review Progress report provides a view of how the documents or items within a case are progressing through the case lifecycle. A typical graph will show a steady progression of files or documents moving through the case. You can use this progression to help predict an end date for the review.

Before you begin: You must have the permission "Allow access to management charts" to view the Case Review Progress report.

To view the progress of a case

1. Select the case. From **Analysis and Review** on the **Dashboard** screen Chart Options box, click the Analysis: drop-down menu and select **Case Review Progress**.



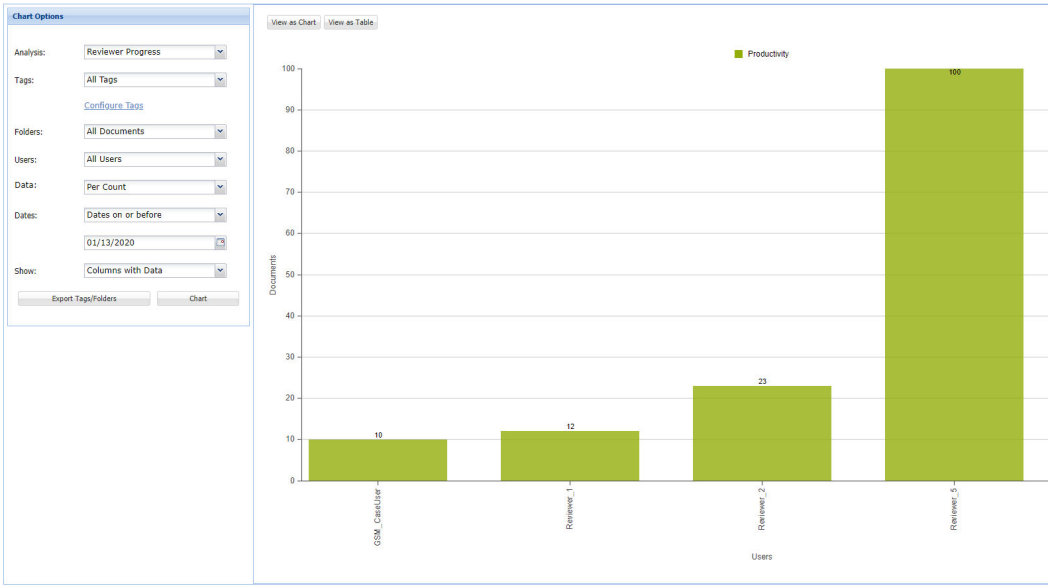
2. Specify the set of criteria to be used to generate the report.
   - All documents, all folders, or a subset of folders
   - All configured tags, a specific tag set, or a specific tag
     - Click the **Configure Tags link** to collect case report statistics for tags by immediately starting a collection job or enable automatic collection to occur periodically.
   - All users or a specified subset of users
   - Documents or items

- Date range
- Time unit desired: daily, weekly, or monthly

3. Click **Chart**.
4. To export this report, or data from all reports, see *"Export Dashboard Reports" on page 216*.

Track Reviewer Progress

The Reviewer Progress report shows how many documents your reviewers have tagged. In order to easily compare reviewer productivity, you can choose to report on how many documents were reviewed per time segment (hour, day, week, or month) or how many documents were reviewed in the date range you are interested in.



Before you begin: You must have the permission "Allow access to management charts" to view the Reviewer Progress report.

To track the progress of a reviewer

1. Select the case. From **Analysis and Review** on the **Dashboard** Chart Options box, click Analysis: **Reviewer Progress**.
2. Specify the set of criteria to be used to generate the report.
   - All documents, all folders, or a subset of folders
   - All tags, a specific tag set, or a specific tag
   - All users or a specified subset of users
   - Documents or items
   - Time unit desired: hourly, daily, weekly, or monthly
   - Date (between, on or before, or on or after) range

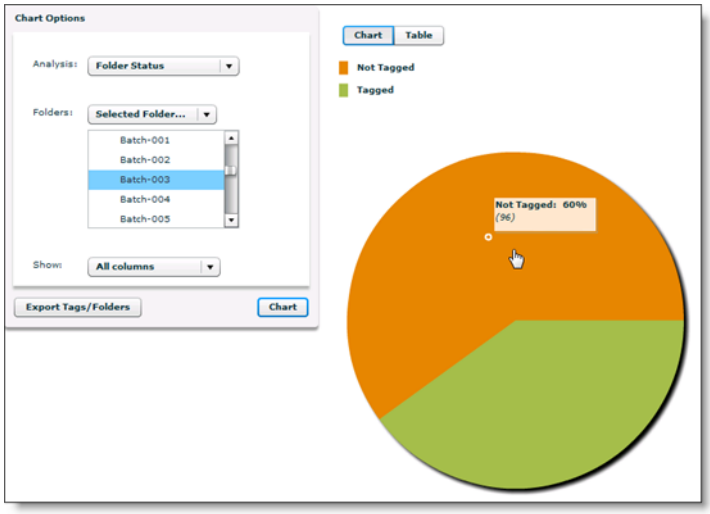
– Show all columns, or columns with data

Note: Hourly productivity statistics are generated by determining the number of documents reviewed while a reviewer is logged in. If the reviewer is logged in and performing non-review tasks (or forgets to log out), that reviewer's hourly productivity statistics can be skewed.

3. Click **Chart**.
4. To view the displayed information in table form, select the **Table** view.
5. To export this report, or data from all reports, see *"Export Dashboard Reports" on page 216*.

View Folder Status

To view the number of tagged or untagged documents within a folder, use the Folder Status report.



To view status of one or more folders

1. Select the case. From **Analysis and Review** in the **Dashboard** Chart Options box, click Analysis: **Folder Status**.
2. From the Folders menu, select all documents, all folders, or a specific folder.

   The "All documents" option displays a pie chart. Charts displaying multiple folders, display bar charts.
3. Select whether you want unused columns to display in the chart.
4. Click **Chart**.

   A chart showing the numbers of documents that are assigned to folders, not assigned to folders, tagged, and not tagged.
5. To view the displayed information in table form, select the **Table** view.
6. To export this report, or data from all reports, see *"Export Dashboard Reports" on page 216*.

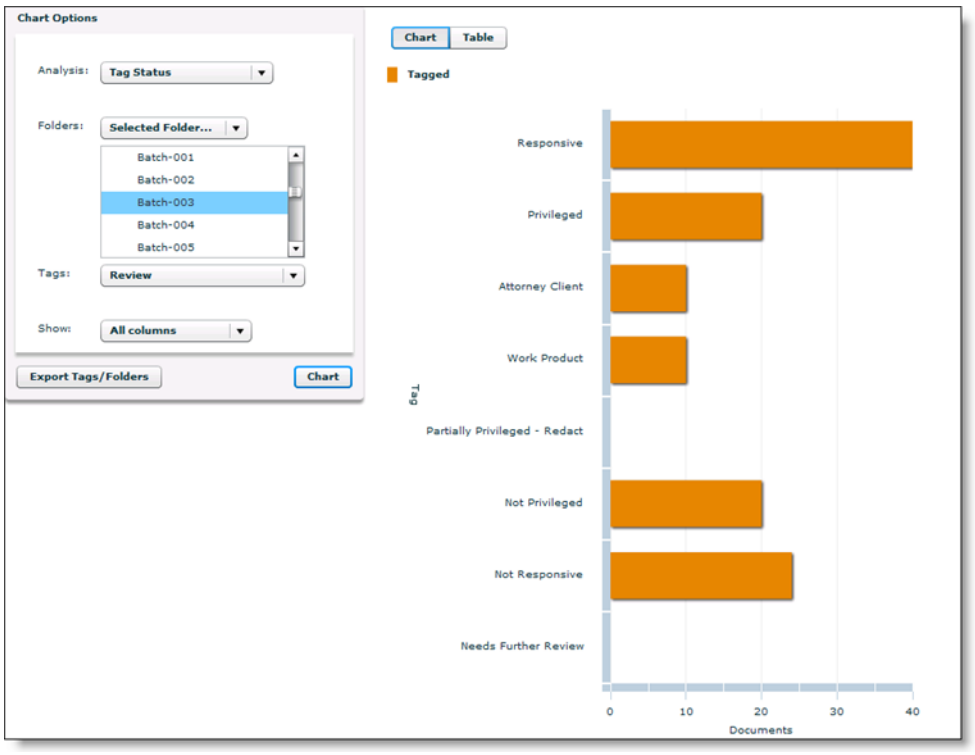
View Tag Status

Before you begin: You must have the permission "Allow analysis tags dashboard access" to view Folder Status reports. The **System Manager**, **Group Admin**, **Case Admin**, **Case Manager**, **Case User**, and **eDiscovery Admin** roles have this permission. Only the **Collections Admin** and **Legal Hold Admin** do not.

To view the status of one or more tags

1. Select the case. From **Analysis and Review** in the **Dashboard** Chart Options box, click Analysis: **Tag Status**.
2. From the Folders menu, select all documents, all folders, or a specific folder.

   All options display bar charts.



3. From the Tags menu, select all tags or a specific tag set.
4. Select whether you want unused columns to display in the chart.
5. Click **Chart**.

   A bar chart displays the number of documents marked with each tag.
6. To view the displayed information in table form, select the **Table** view.
7. To export this report, or data from all reports, see *"Export Dashboard Reports" on page 216*.

View Prediction Rank Statistics

Before you begin: You must have the permission, "Allow analysis tags dashboard access", to view Prediction Rank reports.

To view prediction rank statistics

1. Select the case. From **Analysis and Review** in the **Dashboard** Chart Options box, click Analysis: **Prediction Rank**.
2. From the Folders menu, select all documents, all folders, or a specific folder.
3. From the Model menu, select the predictive tag you want to view.
4. Select the label you want to view.
5. Select whether you want unused columns to display in the chart.
6. Click **Chart**.

   A bar chart displays the number of documents marked with each tag.
7. To view the displayed information in table form, select the **Table** view.
8. To export this report, or data from all reports, see *"Export Dashboard Reports" on this page*.

For more information about predictive coding, see the *Transparent Predictive Coding User Guide.*

Export Dashboard Reports

You can export reports from the dashboard to either a CSV or Excel (XLS) file. Data from a single report, or data from all reports can be exported.

To export dashboard reports to a CSV or Excel (XLS) file

1. Display the report you want to export in table format.
2. Export the report.
   - To export a single report,
     a. Click the Export button to the right.
     b. Select whether you want a CSV file or an XLS file.
     c. **Open** or **Save** the file.
   - To export the data from all of the reports into a single CSV file, click **Export Tags/ Folders**. The export report will appear in the pickup window at the top of the screen: click the compressed file icon to download.

Multiple Language Handling

For information about multiple language support, refer to the following topics:

- *"Language Identification and Best Practices" in the next section*
- *"Multiple Language Search" on page 221*
- *"Frequently Asked Questions" on page 224*
- *"Officially Supported Languages" on page 229*

The application provides multi-language support across the entire eDiscovery platform. The areas supported include:

- Language support

  Supports content in more than 50 languages throughout the complete document life cycle, including processing, analysis, review, and export.

- Stemming support

  Users can enable linguistic stemming for English and more than 10 other languages. Stemmed searches automatically find common variations of search terms, such as "test," "tests," and "testing". Click here for the complete list of languages with stemming support.

- Use of international characters

  Tags, projects, saved searches, and most text input can include international characters.

- Document language identification

  Users can determine how to identify languages based on the amount of that language present in a document. The platform can identify multiple languages in a single document.

- Encoding for export

  Documents are exported in their native encodings. All metadata is exported in a normalized UTF-8 (Unicode) encoding. Non-Unicode encodings such as JIS, Shift-JIS, Big-5, GB are also supported. These are normalized to UTF-8 for indexing, searching, and presentation in the user interface.

All of these functions can be done by users with the **System Manager**, **Group Admin**, **Case Admin**, **Case Manager**, **or eDiscovery Admin** roles.

This section outlines specific details about multi-language functionality that go beyond what is covered in the baseline product documentation. It is targeted at sophisticated users of the platform who need a more detailed understanding of how the application handles various multi-language situations and who may need to be able to do more detailed configuration of cases based on specific multi-language use cases. A list of all supported languages can be found at the end of this document.

Language Identification and Best Practices

Refer to the best practices in this section by understanding how language identification functionality and technology works, particularly when presented with various language challenges and identification settings.

Language Identification Challenges

Historically, language identification has been a challenging problem in eDiscovery applications for two primary reasons:

- False positives/mis-identification

  A mis-identification or false positive occurs when a document is identified as containing a language that it does not have. This is a common challenge for all language detection applications and is exacerbated by the messy and unstructured type of data that is common in electronic discovery. For example, it is very difficult to identify the language of many documents, such as log files, because log files often contain system terms, abbreviations, acronyms, and code words that are hard for language identification algorithms to interpret.

- Noise

  Even if the language in a document is correctly identified, sometimes there is not enough of it to justify special handling. For instance, if an author says something like “c'est la vie” in the middle of an otherwise English-only document, then one would not want to identify this document as containing French.

Language Identification Technology

The application utilizes a language identification engine, Basis Technologies Rosette Linguistics Platform (RLP), and a set of flexible heuristics or rules to perform language identification. Language identification is a challenging problem and different data sets will have different false positive or noise problems. As a result, one set of language identification heuristics may perform well on one data set but not perform well on a different data set.

Multiple Language Document Detection

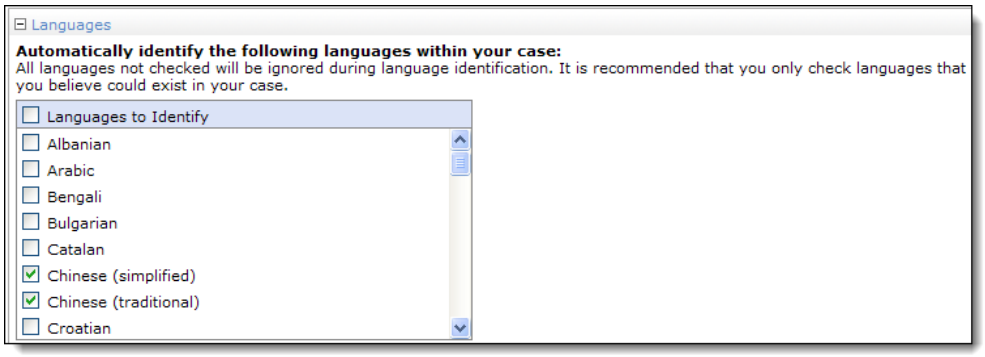
The application supports the ability to detect multiple languages in a document. If, for example, a document contains a substantial amount of both English and Japanese, then the application will identify the document as containing both of these languages.

Language Identification Settings

The application uses a set of user-modifiable heuristics to reduce false positives and noise, allowing you to fine-tune your search results for higher relevance. Users with the correct role can easily modify these settings from the Language section of the Case configuration page under Case Management.

Automatic Language Identification

You can configure the list of languages that will be used for automatic language identification.

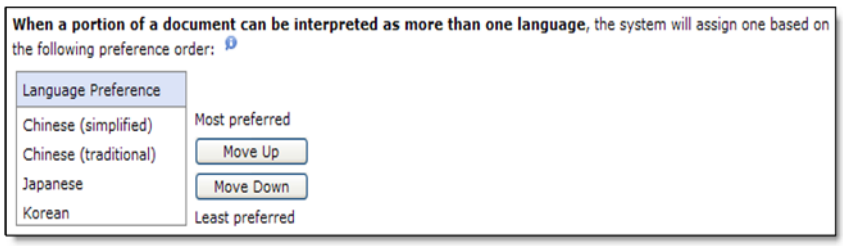


This interface enables you to exclude languages that you know will not exist in your data set and focus only on the ones that you believe could exist. For example, if your client's organization has offices or partners in five countries, then you might decide to only enable automatic identification for languages in these five countries. This not only speeds up document processing, but also reduces the mis-identification of hard-to-identify documents such as log files and Excel spreadsheets. By default, a new case is configured to detect the most common languages including Chinese, English, French, German, Italian, Japanese, Korean, Portuguese, Russian, and Spanish.

Setting Language Preferences

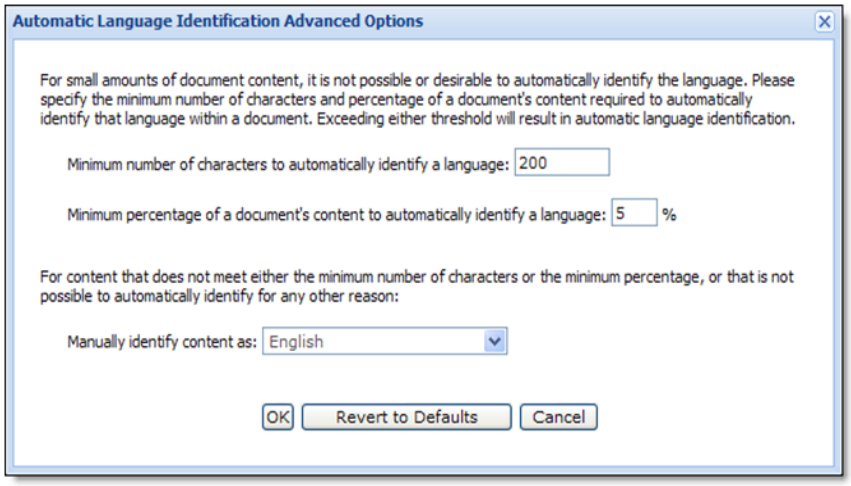
For languages such as Chinese, Japanese and Korean that share common characters, the application allows users to easily set a preference order to specify which language is more likely to occur in their data set. This preference list is used by the application in case there is ambiguity.

Note: The language preference setting cannot be changed after processing has begun.



Setting Advanced Language Options

The application provides a number of Advanced Language options to help you reduce noise. These options allow you to define the amount of content you think is enough to justify special handling.



- Character and Percent Noise thresholds

  You can reduce false positives by specifying the minimum number of characters or the percentage of a document's content required to automatically identify that language within a document. A document will be automatically identified as containing a language if that document contains either the minimum number of characters or the minimum percentage of characters in a language.

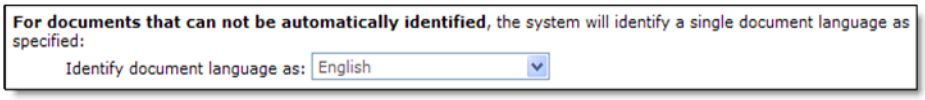
- Noise default language

  If the content does not exceed either the character or percent thresholds specified above, you can specify a language to manually identify it. It may be valuable to use “Other” for this setting if you want to be made aware of language sections that could not be classified. However, if you want to reduce the amount of “noise” in language identification, you can specify a default language to use for unidentified bits. For example, if you know that the vast majority of the content in a case is English, you could choose “English” to automatically classify small bits of content as “English.” Significant chunks of other languages will still be identified per the rules you specify.

As an example, according to the parameters shown in the screenshot above, if an otherwise English document contains the phrase “bon appétit”, it'll still be considered 100% English, if the phrase “bon appétit” doesn't exceed the character or percentage thresholds you have specified for the document. This is done to avoid having the document being assigned to a French linguist when the low occurrence of French doesn't warrant such a review. If the user however wants to get small occurrences of languages reviewed, then we recommend you reduce the character and percentage thresholds.

Setting the Default Case Language

Users can specify a language to use for identifying content that the application does not identify automatically, such as content that could otherwise be mis-identified as a false positive as described above.



If you believe that the majority of your content will be in a particular language, it is recommended that you specify that language with this setting. For US based customers, English will likely be a good choice. You can also choose to use "other" within this setting, which will allow you to segregate documents that were not automatically identified. This is preferred if you want to be notified of documents which could not have their language identified, and review them manually.

Best Practices

Chinese, Japanese, and Korean cases

For cases containing significant amounts of a language that uses characters instead or in addition to a Romanized alphabet, it is best practice to alter default language settings. The reason for this is that Romanized languages use many more characters in a word than character-based languages, such as Chinese, Japanese or Korean. For example, Beijing, China in English, is 12 characters, but only 4 characters in Chinese:

北京中国

When you have a case containing significant amounts of Chinese, Japanese or Korean content, it is suggested that you lower the language thresholds. In particular, change the minimum number of characters threshold in Advanced options from 200 to a lower number appropriate for your data set.

Note: Lowering these thresholds will improve the identification of Chinese, Japanese and Korean but could increase false positives for other languages.

Other cases

Default language configuration values represent the current best practice for non-CJK cases. However, because every data set is different, different values may produce improved results in your particular data set. If improved results are desired, it is recommended to examine the documents in your data set to determine the appropriate language thresholds.

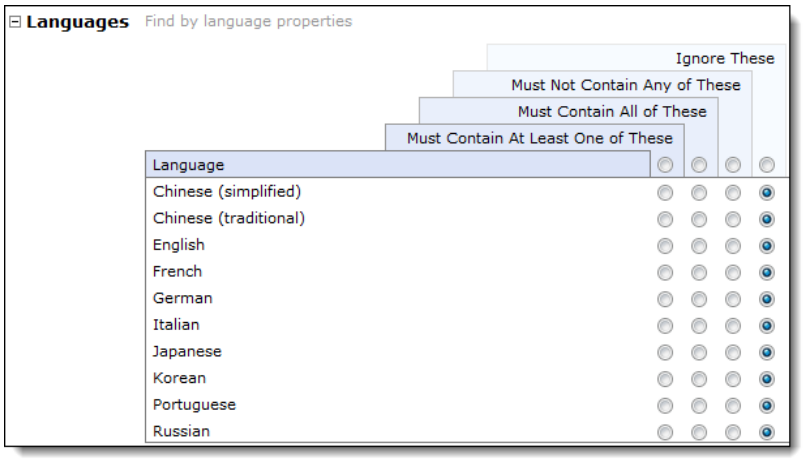
Multiple Language Search

Refer to these key multi-language features and functionality when performing a search on documents containing one or more non-English languages.

Key features

Language Matrix

The application provides a language matrix within advanced search, allowing you to easily select languages for documents that you want to include or exclude from search results.



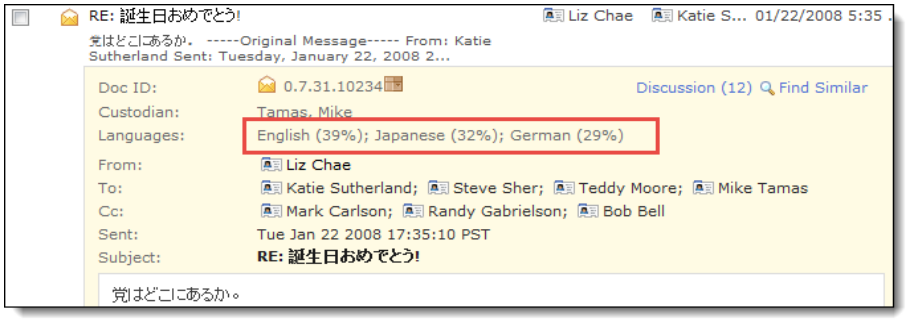
Language Filters

Users can further filter search results by language by checking one or more boxes and clicking Apply Filters. You can filter search results for a particular language with a single click, by clicking on the number next to the language, in parenthesis.

Note: The application can detect multiple languages in a single document, a single document may be counted in one or more of the language filter counts.

Language Composition

For every item in the search results, the language composition which includes a percentage breakdown of each language is displayed.



In the example above, multiple languages are identified and the "English" language at 39% is listed first as the majority language. This language breakdown can be quite useful in complex, l litigation matters involving multiple languages.

Chinese, Japanese, and Korean searches

The application supports searches in all common languages. When performing searches with languages that use characters, such as Chinese, Japanese and Korean, please note the following:

- If you enter characters with no spaces, such as:

  北京中国

  (Beijing China)

  The application will interpret this as a phrase search and will find documents containing these characters in the exact order you specify.

- To search for documents containing ANY of these characters, enter the characters with spaces or using explicit OR operators. For example:

  北京　中国, or 北京 OR 中国

  The application searches for Beijing OR China.

- To search for documents containing ALL of these characters but in no particular order, enter the characters using explicit AND operators. For example:

  北京 AND 中国

  The application searches for Beijing AND China.

Frequently Asked Questions

How many languages does the product support?

We support processing, language identification, search, rendering, and export for over 50 languages including Asian (Chinese, Japanese, Korean), Eastern European (Russian, Bulgarian), Western European (French, German), etc. See Officially Supported Languages in Appendix A for the detailed list.

Is Unicode processed natively?

Unicode only exists to the extent that it is encoded into a byte-level representation of Unicode characters. The application can process Unicode in all of its encodings, as well as process other non-Unicode encodings such as JIS, Shift-JIS, Big-5, GB, and ASCII.

Upon processing, the application represents all Unicode data internally in a UTF-8 encoding across most of its components. On export, all documents are exported in their original encoding), but the metadata contained in the XML output is in UTF-8 format.

What encodings does the product use?

For processing, the application uses a mix of UTF-16 and UTF-8 depending on the document source, type, and stage of processing. Note that many documents that we process may NOT be in Unicode form to begin with; they may be in simple ASCII (single fixed byte) form or some other alternate encoding. For example, Notes documents are stored in the LMBCS ("Lotus Multi-byte Character Set") format. MIME encodings will be converted to Unicode by the application and processed natively. For rendering in the UI and for export, the application uses UTF-8, since this is the best option for displaying content in HTML and representing it in XML.

How does the product handle different encodings?

The application handles encoding conversions through two separate "paths":

- Email processing

  For emails, the application uses MAPI and Notes APIs to take emails from whatever their original encoding is and convert it into Unicode. The application does not directly do any sort of encoding conversion but instead relies completely on the email APIs to do this. This applies to .msg and .eml files and loose file processing.

  Loose files and Email Attachments that have anything other than Windows-1252 encoded text files are passed through Oracle Outside-In. Outside-In has the ability to convert from a set of encodings (shown below) to any other output encoding. We have configured Outside-In to process any of the following encodings on input side, and only Unicode encoded as UTF-8 on output side. Again, the application does not directly do any sort of encoding conversion but relies completely on Outside-In to do the relevant mappings.

Supported Outside-In Character Encodings

Encoding name	Description
iso8859-1	Latin-1
iso8859-2	Latin-2
iso8859-3	Latin-3
iso8859-4	Latin-4
iso8859-5	Cyrillic
iso8859-6	Arabic
iso8859-7	Greek
iso8859-8	Hebrew
iso8859-9	Turkish
macroman	Mac Roman
macce	Mac CE
macgreek	Mac Greek
maccyrillic	Mac Cyrillic
macturkish	Mac Turkish
gb2312	Simplified Chinese
big5	Traditional Chinese
shiftjis	Japanese
eucjp	Japanese
iso2022-jp	Japanese
koi8r	Russian
windows1250	Eastern European
windows1251	Cyrillic
windows1252	Western European
windows1253	Greek
windows1254	Turkish
windows1255	Hebrew
windows1256	Arabic
windows1257	Baltic
thai874	Thai
koreanhangul	Korean Hangul
utf8	UTF-8
unicode	Unicode

Can the product handle Shift-JIS-encoded file names and file paths in container files (such as ZIP/LHZ)?

Normally, file names and file paths are encoded in Unicode. However, in a few cases, the application has identified these to be encoded in Shift-JIS (specifically when data originates from Japan). If you are expecting such data in your case matter, it is recommended that you enable the following property in Support Features. Only users with the **System Manager** role can do this.

To enable container file encoding detection:

1. On the top navigation bar, in System view, click **Support Features**.
2. Select the support feature Property Browser.
3. In the Name of property to change field, type the property:
   `esa.container.filesname.conversion`
4. In the New value field, type:
   true [to enable] false [to disable]
5. Select the check box Confirm change. Are you sure?
6. Click Submit.

For more information on Japanese language encodings and problems with Shift-JIS to Unicode round-tripping, see:

- http://en.wikipedia.org/wiki/Japanese_language_and_computers
- http://web.archive.org/web/20060527013315/http://www.cs.mcgill.ca/~aelias4/encodings.html
- http://support.microsoft.com/kb/170559

There are at least a couple of known instances of encountering an encoding mapping problem due to Japanese characters in filenames. This mis-mapping is referred to as Mojibake (http://en.wikipedia.org/wiki/Mojibake).

Is any client-side software required to use the application on multi-language cases?

No. However, you may need to install fonts (such as Chinese, Japanese, and Korean) if your Windows configuration does not have them installed and enabled by default. If you see characters not being rendered properly in your browser, this should be the first thing that you check because it is the most common reason for the problem.

What CJK characters are actually supported by the product?

As noted above, we only process Unicode in the BMP space. Every CJK Unicode character in the BMP is processed, which is comprised of CJK Unified Ideographs in the range U+4E00 to U+9FFF (20992 characters), and CJK Unified Ideographs Extension A in the range U+3400 to U+4DFF (666 characters), CJK Compatibility Ideographs in the range U+F900 to U+FAFF (512 characters). The following CJK Characters are not supported: Unified Ideographs Extension B U+20000 to U+2A6DF (42720 characters) and CJK Compatibility Ideographs U+2F800 to U+2FA1F (544 characters).

How about stemming support for non-English languages?

Stemming can be enabled as needed for non-English languages during the case setup. The application supports stemming in the following languages:

Stemming and Language Support

Language
Dutch
English (Linguistic and suffix-based)
French
German
Japanese
Korean
Portuguese
Russian

Note: See examples in the next section for the types of stemming that can occur for Japanese.

Can my tags, projects, saved searches, etc. now use international characters?

Yes, almost all user text input may contain international characters.

What encoding format to you use in exporting documents?

Documents are exported out in their native encodings. All metadata is exported in a normalized UTF-8 encoding, the most widely-used and efficient Unicode standard.

Are any 3rd-party components used for multi-language support?

Along with the application's own developed language processing technology, the application Language Processing Engine uses components from Oracle, Basis Technologies (also used by Google), and ICU (International Components for Unicode).

Stemming Examples

Japanese

In general, there are two types of stemming that can occur in Japanese: one for verb conjugation and another for meaning changes on Kanji (Chinese characters within the Japanese language). Following are examples of each type.

Verb Conjugations:

- Example: "To Do":

  します

  The verb "to do" in Japanese is one of the most commonly used, as shown in the these various conjugations that can occur in stemming:

  しました
  する
  した
  される
  された
  されました

- Example: "To Make":

  作ります

  Following are similar verb conjugations that can occur in stemming:

  作りました
  作った
  作って
  作られる
  作られた
  作れる
  作れた

Meaning Changes on Kanji:

- Example: "Tokyo University":

  東京大学

  This name consists of both the word "Tokyo" and "University", which are two separate Kanji groups that are broken up by stemming:

  東京 (Tokyo)　　大学 (University)

- Example: Special Summoning (as in court):

  特殊召喚

  This meaning consists of two Kanji groups that are also broken up by stemming:

  特殊
  召喚

Officially Supported Languages

The following table lists the officially supported languages:

Supported Languages

Language	Officially Supported	Stemming Supported
Albanian	Yes	No
Bengali	Yes	No
Bulgarian	Yes	No
Catalan	Yes	No
Chinese (simplified)	Yes	No
Chinese (traditional)	Yes	No
Croatian	Yes	No
Czech	Yes	No
Danish	Yes	No
Dutch	Yes	Yes
English	Yes	Yes
Estonian	Yes	No
Filipino (Tagalog)	Yes	No
Finnish	Yes	No
French	Yes	Yes
German	Yes	Yes
Greek	Yes	No
Gujarati	Yes	No
Hindi	Yes	No
Hungarian	Yes	No
Icelandic	Yes	No
Indonesian	Yes	No
Italian	Yes	Yes
Japanese	Yes	Yes (See Stemming Example)
Kannada	Yes	No
Korean	Yes	Yes
Kurdish	Transliterated only	No
Latvian	Yes	No
Lithuanian	Yes	No
Malay	Yes	No

Supported Languages

Malayalam	Yes	No
Norwegian (Bokmål)	Yes	No
Pashto	Transliterated only	No
Persian/Farsi	Transliterated only	No
Polish	Yes	No
Portuguese	Yes	Yes
Romanian	Yes	No
Russian	Yes	Yes
Serbian (Cyrillic/Latin)	Yes	No
Slovak	Yes	No
Slovenian	Yes	No
Somali	Yes	No
Spanish	Yes	Yes
Swedish	Yes	No
Tamil	Yes	No
Telugu	Yes	No
Thai	Yes	No
Turkish	Yes	No
Ukrainian	Yes	No
Urdu	Transliterated only	No
Uzbek (Cyrillic/Latin)	Yes	No
Vietnamese	Yes	No

File Types and File Handling

The product supports over four hundred file types and can index text that has been obscured in a number of ways. This section describes some of the file types that the application supports and how obscured text is handled.

For information about file types and file handling, refer to the following topics:

File Types

This section describes special considerations for the file types PST (Microsoft Outlook - Personal Folders File) and NSF (Lotus Notes Mail), MBOX, EMLX. For details on how the application handles these and all other supported file types, see *"File Handling" on page 238*.

PST and NSF Files

The application automatically checks the integrity of PST and NSF files when discovering files within a source to identify potential processing issues. Any PST or NSF file identified with a potential problem is disabled. The file(s) can then be repaired and re-enabled.

Starting with 7.1.3, the product enforces strong file typing for PST and NSF files. An integrity check is performed to identify files that are PST or NSF, but have a different extension such as DOCX. It may be useful to consult the "Not Processed Documents", "Other Type - Extensions" and "Processing Reconciliation" reports to compare file type, file id and file extension information. For more information, see *"Generating Processing Reports" on page 99*.

To minimize the chance of errors, you can perform an initial assessment of your files to identify potential problems. An advanced file search for PST and NSF files within the case file collection can be used to identify files with potential issues.

For PST files, you can run the search and sort the results by file size to identify PSTs that are of a potentially problematic size and then by file attributes to identify read-only PST files.

PST Considerations

The following list represents causes for PST processing errors.

- **Unusual File Sizes.**
  - 2 GB or larger as this is often an indication the PST file is corrupt if it is from Outlook 97-2002.
  - Less than 256 KB in size are often empty or may not be valid PST files.
- **Read-only files.** PST files cannot be read-only. The application requires write access to the PST files to create a write-lock on the file for MAPI access.
- **PST files cannot be password protected.** To provide the public key for decrypting digitally-signed PST messages, go to **System > Support Features** and select the **PKI Certificate Installer** option.
- **Open or in-use files.** PST files cannot be open or in use by Outlook, ScanPST, or any other process while the application is attempting to scan or process them. You should not have Outlook, ScanPST, or any other MAPI tools open on the appliance while processing files or while end-users are accessing the case.

Note: Do not share source files between multiple cases.

NSF Considerations

The following list represents causes for NSF processing errors.

- **Files not shared.** Make sure NSF files are sharable.
- **Access limited or password protected.** Remove any Access Control Lists (ACL) and password protection.

  Note: If you need to remove Access Control Lists from several Lotus Notes files, contact customer support to discuss options for automating this change. See *"Technical Support" on page 11* for more information.
- **Encrypted messages.** NSF files should not have their messages encrypted.
- **Open or in-use files.** NSF files cannot be open or in use by the Lotus Notes client or any other process. You should not have the Lotus Notes client open on the appliance when processing NSF files or while end-users are accessing the case.
- **Truncated messages.** NSF files with truncated messages are automatically disabled by the application during the file integrity check. The administrator can choose to enable these files for processing, but he or she should set the case configuration option correctly to either drop truncated messages, or process them, which makes them available to reviewers with a warning indicator.

OST Files

Starting with version 8.3 CHF2, the eDiscovery platform automates the task of converting OST files to PST files.

OST data files are converted to PST data files during the discovery phase. Upon successful completion of this operation, normal PST processing is applied to the files for easy access, review, and analysis.

Before you begin

System Requirements

Microsoft .NET Framework

You will need a version of Microsoft .NET Framework for the server operating system that is compatible with the OST to PST conversion process. While multiple versions of the .NET Framework can be run on a machine, the OST to PST conversion requires Microsoft .Net Framework 4.5.2 or higher.

New Conversion Libraries Replacing Old Ones

If OST conversion was previously enabled, the upgrade and install process automatically removes older conversion libraries (such as Datanumen) and replaces them with new OST conversion libraries. This means that if OST conversion was in use prior to the upgrade, the conversion service will continue to function but will use the newly installed library

Disk Space

To avoid overwhelming system resources, make sure you have enough disk space for the OST file conversion and any subsequent submissions of partially converted PST files. Key resources, such as available hard disk space can impact overall system functions and will cause partial conversion copies to fail.

Outlook Version Compatibility

OST to PST data file conversion supports OST files from MS Outlook 2010 and 2013.

For MS Outlook 2016, only OST files created from in-place upgrades from Outlook 2010/2013 to MS Outlook 2016 are supported.

Processing partially converted OST files

Starting with 9.0.1, you can configure a property that enables the processing of partially converted OST files. When the ***esa.ost2.conversion_partial_success*** property is set to "**true**" using **System** > **Support Features** > **Property Browser**, the partially converted OST files are processed successfully instead of sending them to the *OST_Partial_ConvertedFiles* directory. By default, the value of the ***esa.ost2.conversion_partial_success*** property is set to "**false**" which results in copying the partially converted OST files to the *OST_Partial_ConvertedFiles* directory. If you choose to use the default settings, you should follow the instructions that are specific to partially converted OST files.

Change the Default Location for Partially Converted OST Files (Optional)

By default, any partially converted OST files will be copied to the system converted files directory:

```
<system converted files directory>\OST_Partial_ConvertedFiles
```

Unless the location was changed by an administrator in system settings, the default for the system converted files directory is:

```
D:\ convertedFiles\OST_Partial_ConvertedFiles
```

You can change this default setting and save files to another location or on another drive by following the instructions below.

To change the location for Partially Converted OST files

1. Go to **System > Support Features**, and select **Property Browser.**
2. Enter the property: `esa.ost2.partialconversion.directory`
3. Set the value to the new location:

   For example:

   ```
   \\\\nassvr.test-t1.local\\data\\case01\\OST_Partial_conversions
   ```

4. Click **Submit** to save the location setting.

Tips for Successful OST File and Partial OST File Conversion

Subdivide Large OST Files

OST files can be quite large. The size depends on many factors including the type of mailbox data (such as email, contact, calendar event, attachments, journal folders). For OST data files that are larger than 2GB, you may want to consider splitting them into smaller parts for faster processing.

OST to PST File Conversion

Make sure you have read and complied with the system and data preparation steps mentioned in the previous section prior to initiating the OST to PST file conversion.To convert OST to PST

To convert OST to PST:

1. Make sure that the Convert supported mailbox files to PST case setting is enabled to take advantage of the OST to PST conversion feature. By default, the new case setting is enabled.
2. During the discovery phase, OST data files are automatically converted into PST data files. No end user input is required.
3. View the job status to view the success/failure/partial success information for the OST job.
4. Identify any errors.

An OST to PST file job might encounter errors. For example, a conversion job can time out before the successful completion. For this type of error and others, check the job log files and follow the steps appropriate for the error that are outlined below.

A. Resolving Incorrect .NET Framework Version

As part of the discovery process, the application checks for a version of the Microsoft .NET Framework that is compatible with the OST conversion tool.
If a missing or incompatible version of .Net Framework is detected, an error message is logged to the discovery job log. The same error appears in the Discovery Errors report.

Error Message Example:

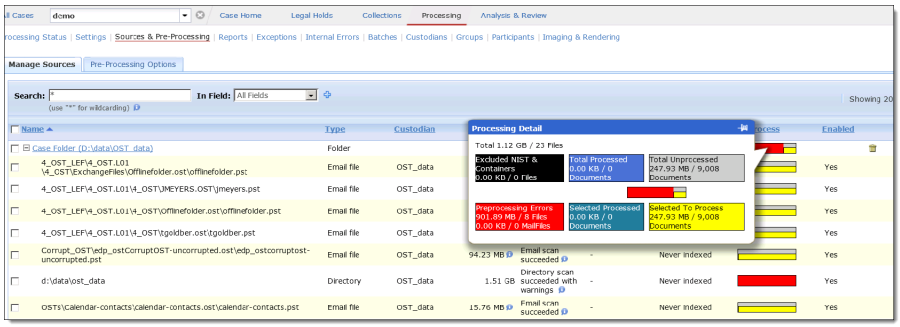
```
ContainerException: ERROR: Discovery of OST file has been
marked as failed because .NET version 4.5.2 or later is
REQUIRED to be installed to successfully convert OST to PST.
Please install the required version of .NET framework
```

The solution involves two steps:

› Install a compatible version of .NET Framework for your server operating system. See System Requirements.

› Copy all the OST files that failed to be discovered to a new folder for rediscovery.

B. Partially Converted Files



Unsuccessful or partially converted OST files display in red as Pre-Processing errors. (Navigate to **Processing > Sources & Pre-Processing>Manage Sources**). Additional details can be obtained by looking at the various logs and the Pre-Processing Error report.

A summary of partially converted files is listed in the job log file. Check the OST2 log under case logs first to find out what type of errors were issued.
Here you will also find a summary of OST files that were partially converted.
To assist with analysis and possible resubmission, the partially converted files are copied to the default system directory or to the directory that you specified in the property browser setting.

Default system directory:

`<system converted files directory>\OST_Partial_Converted-Files`
IMPORTANT: Be sure to examine the Discovery remote job log or Discovery Error report for the exact location and full pathname of all the partially converted OST files. Since the OST_Partial_ConvertedFiles directory contains a fair amount of "noise" that does not apply to partially converted files (such as folder structures with potentially many subdirectories for case ID's and many other elements), it is more efficient to consult the Discovery remote job log and error report for pathname information.

C. Timeout errors

The conversion tool's default settings are calibrated to allow for most OST file conversions to complete. Typically, you do not have to modify any parameter settings. However, there may be instances where these settings may not be sufficient for the conversion profile and data characteristics of extremely large OST files. If this is the situation, you may see conversions of large OST files that terminate with timeout errors. One possible way to resolve the time outs of large OST file conversions is by adjusting limits that are set on the wait time and conversion timeout rate. These settings work together to ensure that the OST to PST conversion workload has sufficient time to process without overwhelming the system resources.
It is important to understand that these settings do not alter or affect the actual OST to PST file conversion rate. This means, for example, that you cannot use these settings to speed up OST conversation rates.
These two settings are accessed from the property browser.

To increase the timeout setting

1. Go to **System > Support Features**, and select **Property Browser**.
2. Enter the property: `esa.OST2.max.timeout.minutes`
3. Set the value to the new timeout value:
   By default, the value is set to 24 hours and the maximum value is 128 hours.
   For example, to enter 100 hours: 6000
4. Click **Submit** to save the timeout setting.

To lower the conversion timeout rate setting

1. Go to **System > Support Features,** and select **Property Browser**.
2. Enter the property: `esa.OST2.mbconvrate.perminute`
3. Set the conversion timeout rate to: 2
4. Click **Submit** to save the timeout setting.
5. Fixing partially converted files
   eDiscovery administrators can choose to process the partially converted PST files into the same case by adding a new case folder that contains the partially converted PSTs and continuing to process the files as a new source.

Note: Partially converted PST files can take up a considerable amount of disk space. Make sure you have deleted any non-essential partially converted PST files to free up disk space.

MBOX

MBOX is an umbrella format which encompasses MBOX, MBX, and a number of other derivative formats that are created by email clients such as Apple Mail, Eudora, and Thunderbird. The application recognizes all of these formats and converts them into PST files prior to processing.

FAQ

Will the application's indexing handle the "blob" format for MBOX messages (e.g. all messages lumped in a single file)?

Yes, the application handles Binary Large Object (BLOB) format.

Do index attachments referenced in a filepath (a la Eudora) vs. encoded in-line in the message, assuming the attachments directory is collected and present?

Yes, the application supports this. However, the attachments directory should be collected in the standard location as for Eudora.

EMLX

EMLX emails are a common Macintosh email format. These are directly supported and do not require conversion.

LEF

Logical Evidence File (LEF) is a container and method for preserving digital evidence which the Veritas eDiscovery platform can process. An LEF can be in L01 format (which is an Encase Forensic proprietary file storage format that stores the file with varying levels of compression and is similar to a ZIP or RAR file).

LEF Considerations

There may be rare instances where the file type contained within a LEF is not recognized. If you encounter such an issues, please contact Technical Support.

File Handling

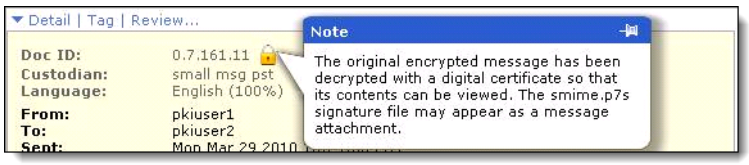
Refer to the following topics in this section:

Encrypted and Digitally-Signed Content

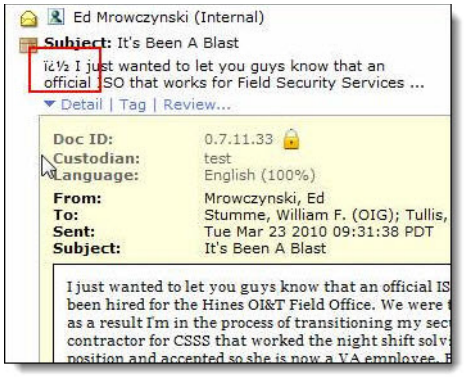
If you provide the public key for decrypting digitally-signed PST messages, the application can now open and index the content.

Encrypted or digitally-signed documents display with a lock icon.

- After uploading PKI digital certificates into the application, you can process, search, analyze and review digitally encrypted or signed PST messages.
- The integrity of original encrypted messages is maintained, while allowing the user to index and view the contents.
- Users can search or filter for encrypted items, and they will display with a lock icon.
- If digital certificates are not installed at the time of processing, the application does not process the messages. These PST files can be reprocessed once the certificates are installed on the server.



Note: To indicate that an item is both PKI encrypted and digitally signed, special characters ï¿½ appear in **Analysis and Review** in front of the message text in Snippet view. These characters are not part of indexed text."



FAQ

Does the application support Encrypted .MSG files?

Currently, only encrypted messages in PST files are supported.

How do I add a certificate to the product?

Upload certificates through the **Support** link from the System menu. For more information, refer to *"Using the Support Features" in the System Administration Guide*.

Once added, does the certificate work forever?

No, if certificate expires it will no longer work.

What happens when I migrate a case to a different appliance?

When you migrate a case with encrypted files to another appliance, you must install the certificates on the new CW machine to view the decrypted files.

All admin accounts on the box should have access to the certificates.

Can I search for encrypted files?

You can search for Encrypted Files in the **Message Flag** section of **Advanced Search**. This returns any files that were originally encrypted (regardless of whether or not the application decrypted them).

Are there any visual clues that a message was encrypted?

A lock icon displays in the header of all documents containing encrypted content.

How does encryption impact my exports?

If the application successfully decrypts a file, then non-native export modes (for example, pdf, html, and so on) will show the decrypted content.

When performing a native export, encrypted documents will be exported as encrypted.

Is processing performance impacted?

Processing performance will clearly slow down when encrypted content is being decrypted and processed.

Hidden Content

By default, the application identifies hidden content within Office and Adobe PDF files. However, this setting can be set/changed to any of the following content handling selections:

- Identify all hidden content
- Extract all documents (for example, non-images)
- Extract images from Notes emails
- Extract images from Office documents
- Extract images from PDF files

  Note: Image extraction from Exchange emails is not currently supported.

For more information, see *"Defining New Cases" on page 18*, and *"Changing the Case Settings" on page 178*. See also the tables in this section for Microsoft Office file handling.

This content includes information that exists within a file but is intentionally or unintentionally obscured so that the user cannot normally view it.

Some of the types of hidden content that the application can identify include:

- Headers and footers: These are generally visible when a document is printed but can occasionally be missed depending on how the document is viewed on the screen
- "Extreme cells": Outlier cells that are contained in hard-to-spot regions of a spreadsheet
- Color obfuscated text: Content such as "white on white" text that isn't visible to someone looking at the document

Starting with 9.1, eDiscovery Platform, by default, attempts to image the following types of documents with hidden content:

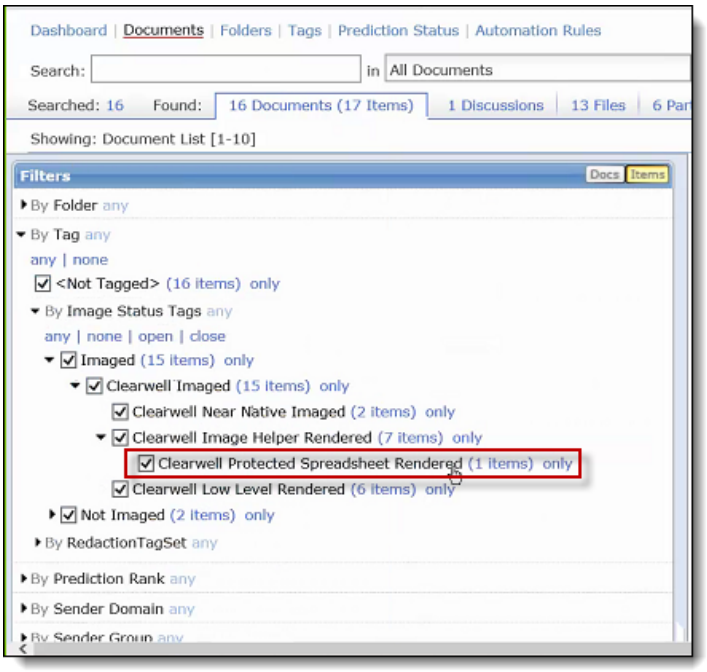
- Tracked changes and comments in Microsoft Word documents
- Hidden rows and columns in Microsoft Excel spreadsheets
- Presenter's Notes in PowerPoint slides

Excel spreadsheets that have been marked as Protected (irrespective of whether they contain hidden content or not) cannot be imaged in the first pass. Such spreadsheets are again sent for imaging after turning off the hidden content extraction.

Protected Excel spreadsheets are indexed and text (including any hidden content) within is made searchable.

Note: Password-protected files are neither indexed nor imaged.

You can use the **Clearwell Protected Spreadsheet Rendered** tag filter to list protected spreadsheets files, which were imaged by the application, but any hidden rows and columns (if present) were excluded in the final image.



Identifying documents containing hidden content

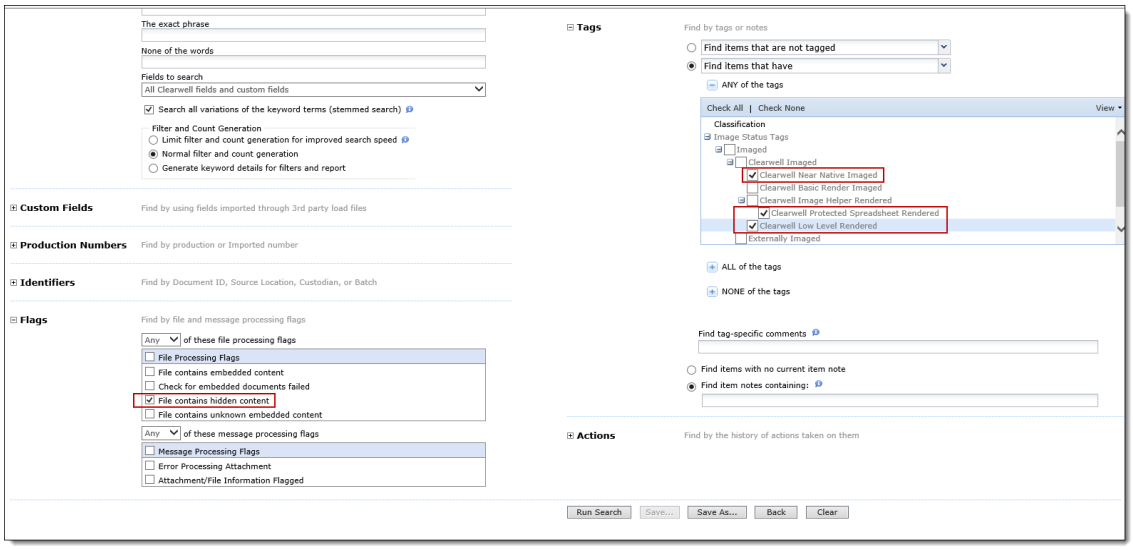
Whenever hidden content is detected within a document, the  icon displays to show the specific type(s) of hidden content that the application was able to identify.

You can identify documents containing hidden content through advanced search and filters. Detailed hidden content information is also included in XML metadata exports.

Identifying imaged items with not-imaged hidden content

You can use advanced search to find successfully imaged items with not-imaged hidden content. If required, you can do a native export on these items and image them using an external imaging tool. You can then perform an import of these images. For detailed steps, see *"Exporting Native Images for External Imaging" on page 94* and *"Import Native Images" on page 94*.

To identify the imaged items with not-imaged hidden content, you should select **File contains hidden content** file processing flag along with **Clearwell Near Native Imaged**, **Clearwell Protected Spreadsheet Rendered**, and **Clearwell Low Level Rendered** Image Status Tags as shown below.



How hidden content is handled

You should consider the following sections that describe how hidden text from Microsoft Excel, Word, and PowerPoint are handled within the application.

Microsoft Excel

Hidden Text	**Indexed**	**Image Output**	**HTML Output**	**Flagged if Contains Hidden Content** (when enabled in Case Settings)
Saved filter setting	No	Print setting and saved filter setting	Yes	No
Text found outside print area	Yes	No	Yes	No
Worksheets	Yes	Yes	Yes	Yes
Rows and columns	Yes	Yes	Yes	Yes
Headers and footers	Yes	Print setting	No	No
White font	Yes	To view text hidden by white font you must change the background color in Native View. **Note:** The TIFF image will not display white text on production export.	To view white text, manually highlight the text. **Note:** Hit highlighting exposes search terms hidden by white font.	Yes
Comments	Yes	No	No	No
Formulas	No	No	No	No
Row and column headings **Note:** The default Excel column headings (A, B, C...) and row headings (1,2,3) are not included.	Yes	Yes	No	No
Tracked changes	No	No	No	Yes

Microsoft Word

Hidden Text	**Indexed**	**Image Output**	**HTML Output**	**Flagged if Contains Hidden Content** (when enabled in Case Settings)
Date fieldcodes	Yes	Yes, however the date displayed is the date the file is rendered.	Yes	No
File fieldcodes	Yes	Yes	Yes	No
Headers and footers	Yes	Yes	No	No
White font	Yes	To view text hidden by white font you must change the background color in Native View. **Note:** The TIFF image will not display white text on production export.	To view white text, manually highlight the text. **Note:** Hit highlighting exposes search terms hidden by white font.	Yes
Comments	Yes	Yes	No	Yes
Tracked changes	Yes, with Clean Content	Yes	Yes	Yes

Microsoft PowerPoint

Hidden Text	**Indexed**	**Image Output**	**HTML Output**	**Flagged if Contains Hidden Content** (when enabled in Case Settings)
Entire slides	Yes	Yes	Yes	Yes
Headers and footers	Yes	Yes	Yes	No
Date and time codes	Yes	Yes, however the date displayed is the date the file is rendered.	No	No

Hidden Text	**Indexed**	**Image Output**	**HTML Output**	**Flagged if Contains Hidden Content** (when enabled in Case Settings)
Comments	Yes, with Clean Content	No	No	Yes
Tracked changes	No	Yes	No	No
Presenter's Notes	No	Yes	No	Yes

Option for not imaging the hidden content

By default, tracked changes and comments in Microsoft Word, hidden rows and columns in Microsoft Excel, and Presenters comments in PowerPoint are attempted for imaging. Though not recommended, you can get the pre-9.1 behavior of not imaging the hidden content in these files by setting the ***esa.muhimbi.unhide.hidden.content*** property value to FALSE by using **System** > **Support Features > Property Browser**.

If you do so, you should consider the following information that describes how hidden text from Microsoft Excel, Word, and PowerPoint are handled within the application.

- When a document is identified with hidden content, you might need to identify the document and view it with the application that created the document.

  For example, hidden content indexed by the tool Clean Content might not *display* in either HTML or Native View. To view the content, open the document in the native application.

- In order to handle dates, headers and footers, the product uses a different tool (not Clean Content) to extract Microsoft Excel hidden content.

Embedded Objects

Embedded objects can be identified and expanded in an "n-level" hierarchy. Once identified and extracted, you can review and redact embedded objects separately from their parents.

Note: As of version 7.0, you can also identify and extract embedded content from RTF, PDF, and all Office documents.

As a part of production, you can also specify whether the embedded objects are produced separately.

FAQ

How do I search for an embedded object within an email?

Embedded objects within emails are always extracted as attachments. To search for an embedded object, go to the Advanced Search page and select **Require attachment or file** from the **File** section.

Why is there no email flag for embedded objects?

Embedded objects within email are displayed as attachments. Within loose files, embedded objects display as embedded objects.

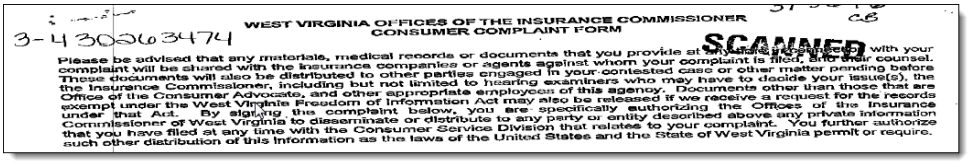
There seem to be two copies of this embedded image. Which version do I redact?

If you need to redact an embedded image, you must redact the image twice: once in the email AND once in the attachment.

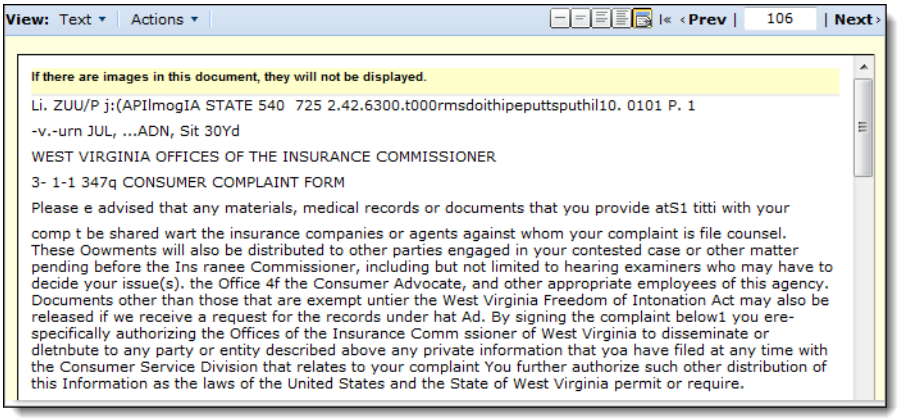
Optical Character Recognition (OCR)

With OCR, the application can read and index content that previously required special handling by reviewers. You can search on and redact text-based content that was previously presented in image file formats.

So, even challenging and hard-to-read content like this:



becomes searchable and reviewable in the application.



Frequently Asked Questions

If you turn on TIFF/txt pairs AND OCR TIFFs, which text gets used? The TIFF/txt text, or the OCR text?

If a TIFF/txt pair is detected, the application excludes the associated TIFF from OCR. The TIFF/txt pair takes priority over OCR.

How does the product handle forms or spreadsheets?

The OCR feature extracts text from forms and tabular information in the correct sequence. For example, images that have text organized in columns will return text inside of a column together (rather than extracting text strictly sequentially left to right, cutting across columns).

In this scenario, proximity searches might not work as expected since extracted text found in one column might not be proximate to text in an adjacent column.

How does the product determine whether a document should be OCR processed?

Documents are processed for OCR based on the rules in this section, and according to the document file extension selected in Case Settings. (For details on OCR settings, see Table in the section Defining New Cases).

The application uses the following logic to determine whether a document needs to be OCR processed.

1. Determine whether document has No Indexed Text.

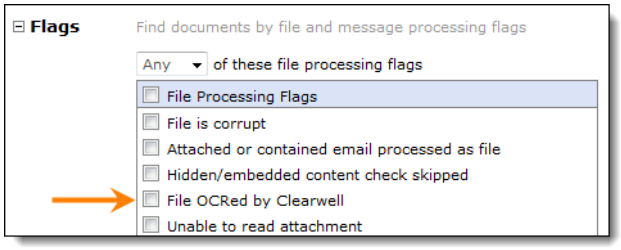
   No Indexed Text means that the document has no text or is below the non-indexed-text threshold. This threshold is set in the Configure OCR Processing section of the **Case > Settings** page.

2. If the document has indexed text or is below the non-indexed text threshold, then do not OCR process the document.

3. If the document does not have indexed text and the size is greater than the OCR threshold, then process the document for OCR.

Can you search for OCR-processed documents?

Yes, you can filter documents by the File Flag, **File OCRed by Clearwell.** You can also search for OCR-processed documents in the Flags section of the Advanced Search page.



Support File Types and File Type Mapping

This section includes a comprehensive reference of all the file formats supported by the application. The file formats are categorized by file type.

Refer to the following file type topics:

Supported Email File Types

Supported File Types (EMail)

Email	Version
Microsoft Outlook PST	Versions 97-2007
Microsoft Outlook MSG	
Microsoft Outlook Express EML	
Lotus Notes NSF (8.x - Windows only, with Domino 8.x Server or Notes 8.x Client - Extraction, conversion, viewing)	Version 6.0 and higher
Apple OS X Mail EMLX	V5.5 and later

Supported Through Conversion to PST

Type	Version
mbox files	V5.5 and later
OST	V8.3 CHF2 and later

Starting with version 8.3 CHF2 and later, the platform converts OST files to PST files. If you are on a pre-8.3 CHF2 version, please refer to this technical article: www.veritas.com/docs/000109381.

Supported Loose File and Email Attachment Types

All Word Processing Documents

Generic Text	Type or Version
ANSI Text	7 and 8 bit
ASCII Text	7 and 8 bit
DOS character set	
EBCDIC	All versions
HTML	Through 4.0 (some limitations)
IBM DCA	
IBM Revisable Form Text	All versions
Macintosh character set	
Microsoft Rich Text Format (RTF)	All versions
Unicode Text	3.0, 4.0
UTF-8	
WML	
XHTML	File ID only
XML	Text only

DOS Word Processors

Name	Type or Version
DEC DX	Through 4.0
DEC DX Plus	4.0, 4.1
Enable	3.0 - 4.5
First Choice	1.0, 3.0
Framework	3.0
IBM DCA/FFT	
IBM DisplayWrite	2.0 – 5.0
IBM Writing Assistant	1.01
Lotus Manuscript	Through 2.0
MASS11	Through 8.0
Microsoft Word	4.0 – 6.0
Microsoft Works	2.0
MultiMate	Through 4.0

DOS Word Processors

Name	Type or Version
MultiMate Advantage	2.0
Navy DIF	All versions
Nota Bene	3.0
Novell PerfectWorks	2.0
Office Writer	4.0 – 6.0
PC-File	5.0
PFS:Write	A, B
Professional Write for DOS	1.0, 2.
QandA Write	2.0, 3.0
Samna Word	Versions through Samna Word IV+
Signature	1.0
SmartWare II	1.02
Sprint	1.0
Total Word	1.2
Wang IWP	Through 2.6
WordMarc	Through Composer Plus
WordPerfect	4.2
WordStar	3.0 – 7.0
WordStar 2000	Through 3.0
XyWrite	Through III Plus

Windows Word Processors

Name	Type or Version
Adobe FrameMaker (MIF)	3.0 – 6.0
Adobe Illustrator Postscript	Level 2
Hangul Version 97, 2002	97 – 2007
JustSystems Ichitaro	5.0, 6.0, 8.0–13.0, 2004
JustWrite	Through 3.0
KingSoft WPS Writer (2010 – Extraction, conversion, viewing)	2010
Legacy	1.1
Lotus AMI/AMI Professional	2.0, 3.0

Windows Word Processors

Name	Type or Version
Lotus WordPro	9.7, 96 – Millennium 9.6
Microsoft Word (2013 – Extraction, conversion, viewing)	98-J, Through 2013
Microsoft WordPad	All versions
Microsoft Works	3.0, 4.0
Microsoft Write	1.0 – 3.0
Novell PerfectWorks	2.0
Novell/Corel WordPerfect (X4 – Extraction, conversion, viewing)	5.1 – X4
OpenOffice Document	
OpenOffice Writer	1.1, 2.0, 3.x
Professional Write Plus	1.0
QandA Write	2.0, 3.0
StarOffice Writer (v9 – Extraction, conversion, viewing)	5.2 – 9
WordStar	1.0

Mac Word Processors

Name	Type or Version
MacWrite II	1.1
Microsoft Word (Mac)	4.0 – 6.0, 98 – 2008
Microsoft Works (Mac)	2.0
Novell WordPerfect	1.02 – 3.1

All Spreadsheets

Name	Type or Version
Enable Spreadsheet	3.0 – 4.5
First Choice SS	Through 3.0
Framework SS	3
IBM Lotus Symphony Spreadsheets	1.x

All Spreadsheets

Name	Type or Version
KingSoft WPS Spreadsheets (Extraction, conversion, viewing)	2010
Lotus 1-2-3	Through Millennium 9.6
Lotus 1-2-3 Charts (DOS and Windows)	Through 5.0
Lotus 1-2-3 for OS/2	2
Microsoft Excel Charts (2013 – Extraction, conversion, viewing)	2.x – 2013
Microsoft Excel for Macintosh	98 – 2008
Microsoft Excel for Windows (2013 – Extraction, conversion, viewing)	3.0 – 2013
Microsoft Excel for Windows (File ID only) (2013 – Extraction, conversion, viewing, no graphics)	2007/2013 Binary
Microsoft Works SS for DOS	2
Microsoft Works SS for Macintosh	2
Microsoft Works SS for Windows	3.0, 4.0
Multiplan	4
Novell PerfectWorks Spreadsheet	2
OpenOffice Calc (3.x – Extraction, conversion, viewing)	1.1 – 3.x
Openoffice Spreadsheet	
PFS: Plan	1
QuattroPro for DOS	Through 5.0
QuattroPro for Windows	Through X3
SmartWare II SS	1.02
SmartWare Spreadsheet	
StarOffice Calc (v9 – Extraction, conversion, viewing)	5.2 – 9
SuperCalc	5
Symphony	Through 2.0
VP	Planner 1
WordPerfect Spreadsheets (X4 – Extraction, conversion, viewing)	X4

All Presentations

Name	Type or Version
Corel Presentations	
Harvard Graphics Presentation DOS	3.0
IBM Lotus Symphony Presentations	1.x
KingSoft WPS Spreadsheets (Extraction, conversion, viewing)	2010
Lotus Freelance	1.0–Millennium 9.6
Lotus Freelance for OS/3	2
Lotus Freelance for Windows	95, 97
Microsoft PowerPoint for Macintosh	4.0 – 2008
Microsoft PowerPoint for Windows (2013 – Extraction, conversion, viewing)	3.0 –2013
Novell Presentations	3.0, 7.0
OpenOffice Impress (3.x – Extraction, conversion, viewing)	1.1, 2.0, 3.x
StarOffice Impress (v9 – Extraction, conversion, viewing)	5.2 – 9
WordPerfect Presentations (X4 – Extraction, conversion, viewing)	6.0 – X4

All Images

Name	Type or Version
Adobe Illustrator	4.0 - 7.0, 9.0
Adobe Illustrator (XMP only)	11 – 13 (CS 1 – 3))
Adobe InDesign (XMP only)	3.0 – 5.0 (CS 1 - 3)
Adobe InDesign Interchange (XMP only)	
Adobe PDF	1.0 – 1.7 (Acrobat 1 - 9)
Adobe PDF Package (Extraction, conversion, viewing)	
Adobe PDF Portfolio (Extraction, conversion, viewing)	
Adobe Photoshop	4.0
Adobe Photoshop (XMP only)	8.0 – 10.0 (CS 1 – 3)

All Images

Name	Type or Version
Ami Draw	SDW
AutoCAD Drawing	2.5, 2.6
AutoCAD Drawing	9.0 – 14.0
AutoCAD Drawing	2000 - 2007
AutoShade Rendering	2
CALS Raster (GP4)	Type I
CALS Raster (GP4)	Type II
Computer Graphics Metafile	ANSI
Computer Graphics Metafile	CALS
Computer Graphics Metafile	NIST
Corel Draw	2.0 – 9.0
Corel Draw Clipart	5.0, 7.0
Encapsulated PostScript (EPS)	TIFF header Only
Enhanced Metafile (EMF)	
Escher graphics	
FrameMaker Graphics (FMV)	3.0 – 5.0
Gem File (Vector)	
GEM Image (Bitmap)	
Graphics Interchange Format (GIF)	
Harvard Graphics Chart DOS	2.0 – 3.0
Harvard Graphics for Windows	
HP Graphics Language	2.0
IBM Graphics Data Format (GDF)	1.0
IBM Picture Interchange Format	1.0 s
IGES Drawing	5.1 – 5.3
JBIG2	Graphic Embeddings in PDF
JFIF (JPEG not in TIFF format)	
JPEG	
JPEG 2000	JP2
Kodak Flash Pix	
Kodak Photo CD	1.0
Lotus PIC	
Lotus Snapshot	

All Images

Name	Type or Version
Macintosh PIC	BMP only
Macintosh PICT2	BMP only
MacPaint	
Micrografx Designer	Through 3.1
Micrografx Designer	6.0
Micrografx Draw	Through 4.0
Microsoft Windows Bitmap	
Microsoft Windows Cursor	
Microsoft Windows Icon	
Microsoft XPS (Text only)	
Novell PerfectWorks Draw	2
OpenOffice Draw	1.1 – 3.x
OS/2 Bitmap	
OS/2 Warp Bitmap	
Paint Shop Pro (Win32 only)	5.0, 6.0
PC Paintbrush (PCX)	
PC Paintbrush DCX (multi-page PCX)	
Portable Bitmap (PBM)	
Portable Graymap PGM	
Portable Network Graphics (PNG)	
Portable Pixmap (PPM)	
Progressive JPEG	
StarOffice Draw (v9 – Extraction, conversion, viewing)	6.x – 9
Sun Raster	
TIFF	Group 5 and 6
TIFF CCITT	Group 3 and 4
TruVision TGA (Targa)	2.0
Visio	5.0 - 2007
Visio (Page Preview mode WMF/EMF)	4.0
Visio XML VSX (File ID only)	2007
WBMP wireless graphics format	
Windows Metafile	

All Images

Name	Type or Version
Word Perfect Graphics (X4 – Extraction, conversion, viewing)	1.0, 2.0 – 10.0, X4
X-Windows Bitmap	x10 compatible
X-Windows Dump	x10 compatible
X-Windows Pixmap	x10 compatible

Database

Name	Type or Version
DataEase	4.x
DBase	III, IV, V
First Choice DB	Through 3.0
Framework DB	3.0
Microsoft Access	1.0, 2.0
Microsoft Works DB for DOS	2.0
Microsoft Works DB for DOS	1.0
Microsoft Works DB for Macintosh	1.0
Microsoft Works DB for Windows	3.0, 4.0
Paradox for DOS	2.0 – 4.0
Paradox for Windows	2.0 – 4.0
QandA Database	Through 2.0
R:Base	R:Base 5000
R:Base	R:Base System V
Reflex	2.0
SmartWare II DB	1.02

All Multimedia (Sound and Video)

Name	Type or Version
AVI (Metadata extraction only)	
Flash (text extraction only)	6.x, 7.x, Lite
MP3 (ID3 metadata only)	

All Multimedia (Sound and Video)

Name	Type or Version
MPEG-1 Audio layer 3 V ID3 v1 (File ID only)	
MPEG-1 Audio layer 3 V ID3 v2 (File ID only)	
MPEG-1 Video V 2 (File ID only)	
MPEG-1 Video V 3 (File ID only)	
MPEG-2 Audio (File ID only)	
MPEG-4 (Metadata extraction only)	
MPEG-7 (Metadata extraction only)	
QuickTime (Metadata extraction only)	
Real Media - (File ID only)	
WAV (Metadata extraction only)	
Windows Media ASF (Metadata extraction only)	
Windows Media Audio WMA (Metadata extraction only)	
Windows Media DVR-MS (Metadata extraction only)	
Windows Media Video WMV (Metadata extraction only)	
All Multimedia (Sound and Video)	

Other Types

Name	Type or Version
Microsoft Project (File ID only)	2007
Microsoft Project (text only)	98 – 2003
Microsoft Windows DLL	
Microsoft Windows Executable	
vCalendar	2.1
vCard	2.1
Yahoo! Messenger	6.x – 8

Supported Container Extraction File Types

By default, the application will extract the contents of the following container file types during processing. Container extraction can be disabled during case setup.

Supported Container Extraction File Types

Name	Type or Version	File ID	Cannot Exclude
ZIP		1802,	
RAR		1821	
TAR		1807	
LZH (and LHA)		1813, 1814	
JAR		1802	
GZIP		1815	
Self-Extracting ZIP files (.exes)		1803	
Self-Extracting RAR files (.exes)		1822	
UNIX_COMP		1806	
BZ2 (bzip2)		65537	
7Zip		65538,1826, 1827	
LEF (.L01)	V5.1 and later		✓
E01	V5.1 and later		✓
MBOX		1817	✓
OST	V8.3 CHF2 or later		

Detection of supported container files is performed by looking at the actual file content, not simply by file extension. As a result, it is possible that additional formats are also supported because they are in fact identical to the officially supported formats. For example, DEB and AR files are usually similar enough to TAR that they can be extracted.

If an unsupported container format is encountered, it will be treated as a loose file/attachment during processing.

Note: Container File IDs may be useful for the "Not Processed Documents", "Other Type - Extensions", "Processing Reconciliation" reports. See *"Generating Processing Reports" on page 99*.

File Type Mapping

File Type Mapping

File Type	Mapping
Adobe Acrobat	PDF, PDFIMAGE
Microsoft Word	WORD4, WORD5, MACWORD3, MACWORD4, WINWORD1, WINWORD1COMPLEX, WINWORD2, MACWORD5, WORD6, WINWORD6, WINWORD1J, WINWORD5J, WINWORD2_OLECONV, WINWORD7, MACWORD6, WINWORD97, MACWORD97, WINWORD2000, WINWORD2002, WINWORD2003, WORDXML12, WINWORD2007, ENCRYPTED_WORD2007, WINWORDTEMPLATE2007, DRM_WORD, DRM_WORD2007
Microsoft Excel	EXCEL, EXCEL3, EXCEL4, EXCEL5, MACEXCEL4, MACEXCEL5, EXCEL97, EXCEL3WORKBOOK, EXCEL4WORKBOOK, MACEXCEL4WORKBOOK, REGMACEXCEL4WB, EXCEL2000, EXCEL2002, EXCEL2003, EXCEL2007, ENCRYPTED_EXCEL2007, EXCEL2007_BINARY, DRM_EXCEL, SSEND
Microsoft Power Point	POWERPOINT4, POWERPOINT3, POWERPOINT7, POWERPOINTMAC3, POWERPOINTMAC4, EXTPOWERPOINT4, EXTPOWERPOINTMAC4, POWERPOINTMACB3, POWERPOINTMACB4, POWERPOINT97, POWERPOINT9597, POWERPOINT2000, POWERPOINT2, POWERPOINT2007, ENCRYPTED_PPT2007, DRM_POWERPOINT, DRM_POWERPOINT2007
Email (.eml file)	MIMEOUTLOOKEML, TEXTMAIL, MIMEMAIL, EMLX
Email (.msg file)	OUTLOOK_MSG

File Type Mapping

File Type	Mapping
All word processing documents	WORD4, WORD5, WORDSTAR5, WORDSTAR4, WORDSTAR2000, WORDPERFECT5, MULTIMATE36, MULTIMATEADV, RFT, TXT, SMART, SAMNA, PFSWRITEA, PFSWRITEB, PROWRITE1, PROWRITE2, IBMWRITING, FIRSTCHOICE, WORDMARC, DIF, VOLKSWRITER, DX, SPRINT, WORDPERFECT42, TOTALWORD, IWP, WORDSTAR55, WANGWPS, RTF, MACWORD3, MACWORD4, MASS11PC, MACWRITEII, XYWRITE, FFT, MACWORDPERFECT, DISPLAYWRITE4, MASS11VAX, WORDPERFECT51, MULTIMATE40, QAWRITE, MULTIMATENOTE, PCFILELETTER, MANUSCRIPT1, MANUSCRIPT2, ENABLEWP, WINWRITE, WORKS1, WORKS2, WORDSTAR6, OFFICEWRITER, MACWORD4COMPLEX, DISPLAYWRITE5, WINWORD1, WINWORD1COMPLEX, AMI, AMIPRO, FIRSTCHOICE3, MACWORDPERFECT2, MACWORKSWP2, PROWRITEPLUS, LEGACY, SIGNATURE, WINWORDSTAR, WINWORD2, JUSTWRITE, WORDSTAR7, WINWORKSWP, JUSTWRITE2, AMICLIP, LEGACYCLIP, PROWRITEPLUSCLIP, MACWORD5, ENABLEWP4, WORDPERFECT6, WORD6, DX31, WPFENCRYPT, QAWRITE3, MACWORDPERFECT3, CEOWORD, WINWORD6, WORDPERFECT51J, ICHITARO3, ICHITARO4, WINWORD1J, WINWORD5J, MATSU4, MATSU5, P1, RTFJ, CEOWRITE, WINWORKSWP3, WORDPAD, WPFUNKNOWN, WINWORD2_OLECONV, WORDPERFECT61, FTDF, WORDPERFECT5E, WORDPERFECT6E, HTML, WINWORD7, AREHANGEUL, HANA, WINWORKSWP4, PERFECTWORKS1, WORDPERFECT7, WORDPRO, HTML_LATIN2, HTML_JAPANESESJIS, HTML_JAPANESEEUC, HTML_CHINESEBIG5, HTML_CHINESEEUC, HTML_CHINESEGB, HTML_KOREANHANGUL, HTML_CYRILLIC1251, HTML_CYRILLICKOI8,CYRILLIC1251, CYRILLICKOI8, WWRITE_SHIFTJIS, WWRITE_CHINESEGB, WWRITE_HANGEUL, WWRITE_CHINESEBIG5, WPSPLUS, MACWORD6, WINWORD97, RAINBOW, INTERLEAF, MACWORD97, INTERLEAFJ, WORDPERFECT8, ICHITARO8, VCARD, HTML_CSS, POCKETWORD, WORDPRO97, WINWORD2000, W2KHTML, XL2KHTML, PP2KHTML, XML, WML, WMLB, HTML_JAPANESEJIS, WML_CHINESEBIG5, WML_CHINESEEUC, WML_CHINESEGB, WML_CYRILLIC1251, WML_CYRILLICKOI8, WML_JAPANESEJIS, WML_JAPANESESJIS, WML_JAPANESEEUC, WML_KOREANHANGUL, WML_LATIN2, WML_CSS, STAROFFICEWRITER52, MIFF6, MIFF6J, MIFF, JAVASCRIPT, TEXT, HDML, CHTML, XHTMLB, HTMLAG, HTMLWCA, SEARCHML, POCKETWORD20, WIRELESSHTML, HANGULWP97, HANGULWP2002, HTMLUNICODE, XML_DOCTYPE_HTML, PAGEML, EBCDIC, WINWORD2002, WINWORD2003, MIME, STAROFFICEWRITER6, OUTLOOK_PST, XHTML, MSWORKS2000, MIMENEWS, MIMEOUTLOOKNEWS, VCAL, TNEF, MHTML, WPEND, SMARTDATA, FRAMEWORKIII, WORKSDATA, DATAEASE, MSPROJECT98, MSPROJECT2000, SEARCHTEXT, PSTF, PST_2003, PAB_2002, SEARCHML20, SEARCHML30, YAHOOIM, WORDXML2003, WORDXML12, STAROFFICEWRITER8, SEARCHML31, OUTLOOK_OFT, WINWORD2007, ENCRYPTED_WORD2007, WINWORDTEMPLATE2007, SEARCHML32, DRM_UNKNOWN, DRM_WORD, DRM_WORD2007

File Type Mapping

File Type	Mapping
All spreadsheets	SYMPHONY1, 123R1, 123R2, 123R3, SMARTSHEET, EXCEL, ENABLESHEET, WORKSSHEET, VPPLANNER, TWIN, SUPERCALC5, QUATTROPRO, QUATTRO, PFS_PLAN, FIRSTCHOICE_SS, EXCEL3, GENERIC_WKS, MACWORKSSS2, WINWORKSSS, EXCEL4, QUATTROPROWIN, 123R4, QUATTROPRO1J, CEOSS, EXCEL5, MULTIPLAN4, WINWORKSSS3, QUATTROPRO4, QUATTROPRO5, QUATTROPRO6, 123R2OS2, 123R2OS2CHART, WINWORKSSS4, QUATTROPRO7NB, QUATTROPRO7GR, 123R6, MACEXCEL4, MACEXCEL5, EXCEL97, EXCEL3WORKBOOK, EXCEL4WORKBOOK, MACEXCEL4WORKBOOK, REGMACEXCEL4WB, 123R9, QUATTROPRO8, QUATTROPRO9NB, EXCEL2000, QUATTROPRO10NB, EXCEL2002, STAROFFICECALC52, QUATTROPRO11NB, EXCEL2003, STAROFFICECALC6, QUATTROPRO12NB, STAROFFICECALC8, EXCEL2007, EXCEL2007_BINARY, DRM_EXCEL, SSEND
All images	BMP, TIFF, PCX, GIF, EPSTIFF, CCITTGRP3, MACPICT2, WPG, WINDOWSMETA, LOTUSPIC, MACPICT1, AMIDRAW, TARGA, GEMIMG, OS2DIB, WINDOWSICON, WINDOWSCURSOR, MICROGRAFX, MACPAINT, WPG2, CGM, CANDY4, HANAKO1, HANAKO2, JPEGFIF, DCX, OS2METAFILE, DXFA, DXFB, DXB, OS2WARPBMP, WPG7, SUNRASTER, KODAKPCD, ENHWINDOWSMETA, GEM, IGES, IBMPIF, XBITMAP, XPIXMAP, CALSRASTER, PNG, XDUMP, GDF, DESIGNER, PBM, PGM, PPM, ADOBEPHOTOSHOP, PAINTSHOPPRO, FLASHPIX, PROGRESSIVEJPEG, DGN, BMP5, WBMP, MIFFG, WPG10, EXPORTIMAGE, OS2V2BMP
All multimedia (sound and video)	RIFFWAVE, RIFFAVI, MIDI, DIRECTOR, FLASH6, QUICKTIME, MP3_ID31, MP3_ID32, ID31, ID32, MP3, MPGAV1L1, MPGAV1L2, MPGAV2L1, MPGAV2L2, MPGAV2L3, ASF, WMV, WMA, DVR_MS, REALMEDIA, MPEG1, MPEG2, ISOBASEMEDIAFILE, MPEG4, MULTIMEND
All programs	EXECUTABLE, COM, ZIPEXE, MSCAB
Other types	(file types not found above)

Appendix A: Product Documentation

The table below lists the administrator and end-user documentation that is available for the Veritas eDiscovery Platform product.

Veritas eDiscovery Platform Documentation

Document	Comments
Installation and Configuration	
Installation Guide	Describes prerequisites, and how to perform a full install of the Veritas eDiscovery Platform application
Upgrade Overview Guide	Provides critical upgrade information, by version, useful prior to upgrading an appliance to the current product release
Upgrade Guide	Describes prerequisites and upgrade information for the current customers with a previous version of the software application
Utility Node Guide	For customers using utility nodes, describes how to install and configure appliances as utility nodes for use with an existing software setup
Distributed Architecture Deployment Guide	Provides installation and configuration information for the Review and Processing Scalability feature in a distributed architecture deployment
Getting Started	
Navigation Reference Card	Provides a mapping of review changes from 10.x compared to 9.x, 8.x compared to 7.x, and the user interface changes from 7.x compared to 6.x
Administrator's QuickStart Guide	Describes basic appliance and case configuration
Reviewer's QuickStart Guide	A reviewer's reference to using the Analysis and Review module
Tagging Reference Card	Describes how tag sets and filter type impact filter counts
User and Administration	
Legal Hold User Guide	Describes how to set up and configure appliance for Legal Holds, and use the Legal Hold module as an administrator
Identification and Collection Guide	Describes how to prepare and collect data for processing, using the Identification and Collection module
Case Administration Guide	Describes case setup, processing, and management, plus pre-processing navigation, tips, and recommendations. Includes processing exceptions reference and associated reports, plus file handling information for multiple languages, and supported file types and file type mapping
System Administration Guide	Includes system backup, restore, and support features, configuration, and anti-virus scanning guidelines for use with Veritas eDiscovery Platform
Load File Import Guide	Describes how to import load file sources into Veritas eDiscovery Platform
User Guide	Describes how to perform searches, analysis, and review, including detailed information and syntax examples for performing advanced searches

Veritas eDiscovery Platform Documentation

Document	Comments
Imaging Tool Upgrade Guide	Provides details about the Imaging Tool Upgrade feature and how to perform Imaging Tool Upgrade after the eDiscovery Platform appliance is upgraded to version 10.0, workflows affected when the cases are upgraded or not upgraded, and frequently asked questions (FAQs).
Export and Production Guide	Describes how to use and produce exports, productions, and logs (privilege and redaction logs)
Transparent Predictive Coding User Guide	Describes how to use the Transparent Predictive Coding feature to train the system to predict results from control data and tag settings
Audio Search Guide	Describes how to use the Audio Search feature to process, analyze, search and export search media content
Reference and Support	
Audio Processing	A quick reference card for processing multimedia sources
Audio Search	A quick reference card for performing multimedia search tasks
Legal Hold	A quick reference card of how to create and manage holds and notifications
Collection	A quick reference card of how to collect data
OnSite Collection	A quick reference for performing OnSite collection tasks
Review and Redaction	Reviewer's reference card of all redaction functions
Keyboard Shortcuts	A quick reference card listing all supported shortcuts
Production	Administrator's reference card for production exports
User Rights Management	A quick reference card for managing user accounts
Online Help	
Includes all the above documentation (excluding Installation and Configuration) to enable search across all topics. To access this information from within the user interface, click **Help**.	
Release	
Release Notes	Provides latest updated information specific to the current product release