Harvester Server

Pinpoint Labs Harvester Server is a software application that enables users to collect data on remote systems from one shared location.

Users are able to create individual projects for different collection needs, launch and manage jobs from a remote location.

Users are also able to create, edit and manage jobs that can be used for any project.

Harvester Server can be tailored to suit remote collection needs.

Activating Harvester Server

ONLINE ACTIVATION

Harvester Server can be licensed and run from a shared network location or host computer.

Reference Video: [Harvester Server 5.1 Download and Activate]

To install and activate, follow these steps:

- Download the Harvester ZIP file specified in your registration email.
- Unzip the entire Harvester ZIP file contents to a shared folder or drive.
- Double click the Harvester Server application File ***Harvester Server.exe*** via its UNC path.
- When the new window displays asking "Do you want to register Harvester now?" click Yes

- Enter Account ID and click ***Register*** to activate the product using online activation.

OFFLINE REGISTRATION

Should a firewall block the connection to the licensing server, offline activation can be accomplished by doing the following:

- Enter your account ID and click **Offline Registration.**
- A serial number will be generated; call 1-888-304-1096 or email *support@pinpointlabs.com* with your Account ID and the serial number.
- Pinpoint Labs will generate an Activation key for you; enter this in the space provided and click **Register**.

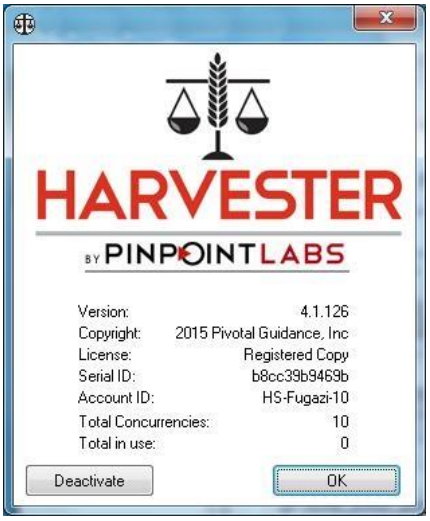
TRANSFER LICENSE

Another advantage of Harvester Server is the ability to move it from location to location. Once activated on a server, you can perform your collection, deactivate and move the license to another network or server. The **Tools** menu has an option to **Deactivate License**, restoring the license to be used again elsewhere. Deactivating the license requires internet access.

Once you have placed the .zip file contents onto another device or computer, you can activate the license using your same Account ID.

ACCESSING PRODUCT AND LICENSING DETAILS

The items listed below can be accessed in the Harvester Server **About** screen from the **Registration** icon located under the Tools Tab as shown below:



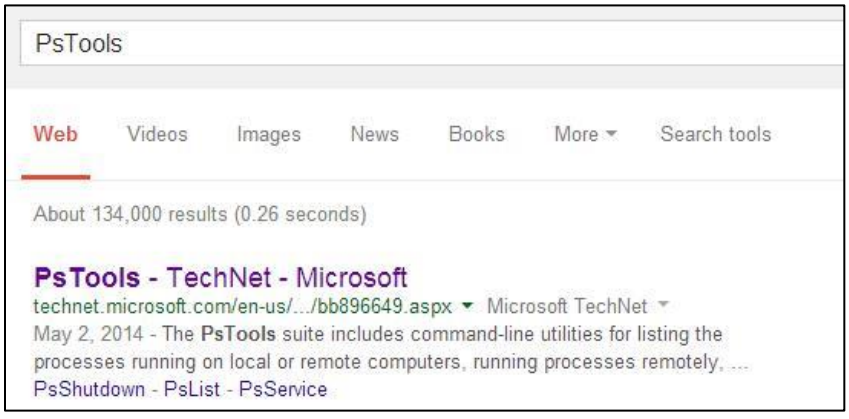
- Version
- Serial ID
- Account ID
- Total Concurrences
- Deactivate (License)
  NOTE: If using Harvester Hybrid licenses, **Manage Licenses** will also appear in the About screen

Deactivating the server license will disable all concurrences. Once the server license is successfully deactivated, Harvester can be registered on a different location.

ADDITIONAL FILE REQUIREMENTS

To be able to launch jobs remotely via PsExec or Relay Auto Deploy, you will need to download and add additional files to the Harvester Server Folder. To download and add these files, follow these steps:

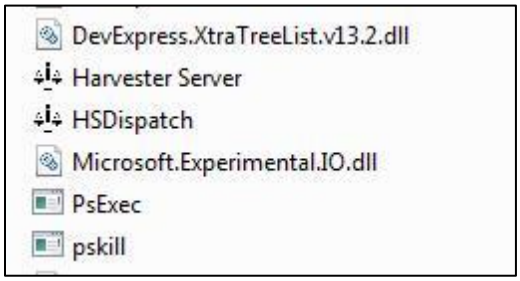
- Open the internet browser of your choice and search for PsTools.



- Download the *PsTools Suite* from Microsoft.com

NOTE: This is a free download

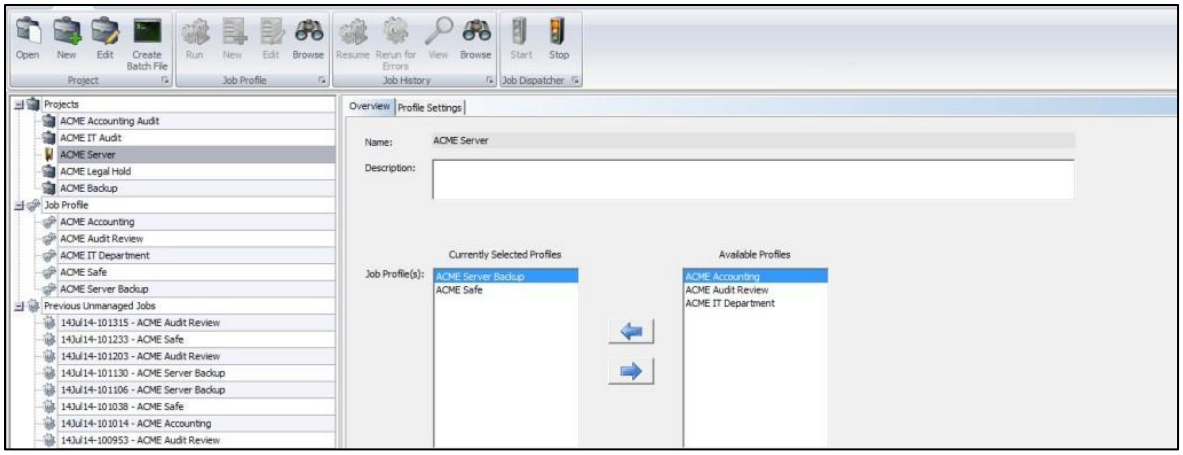
- Once the PsTools suite is downloaded, extract the files and select the ***PsExec*** and ***pskill*** application files. Drag and drop them into the Harvester Server folder.



Creating a Project

Reference Video: **Harvester Server Advanced Part 1- Project Queue and Dispatch Submission Options**

Projects simplify organization and distinguish between different Harvester Server collection tasks. Upon starting Harvester Sever for the first time, the project list will be empty.



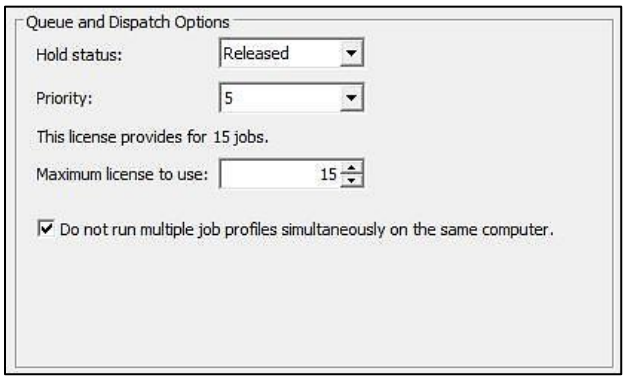
Create your first project with Harvester Server.

- Select **New** in the project section of Harvester Server; this will open a clean template for creating a project.
- Fill in a name that can be used to separate this project from other projects that may be in the project tree. This name will be the permanent project title.
- Select any jobs to be *selected by default* in the Job Queue Manager. The profiles can be edited in the Job Queue Manager before launching.

***NOTE*:** All profiles will be available in the Job Queue Manager; however, profiles selected here will be selected by default.

PROFILE SETTINGS TAB

QUEUE AND DISPATCH OPTIONS



***Hold Status*:** The hold status will set the default of each job submitted within Harvester Server to Held or Released. This will determine whether the administrator will have to manually

release each job or whether they will be released automatically.
***Priority*:** The priority set by the user will determine if the project will launch submitted collection jobs compared to other projects that have submitted collection jobs. Higher priority projects will launch jobs before lower priority projects.
Maximum licenses to use: This option allows the user to select how many of the available licenses they want the project to use. For example, if the user has Harvester Server with three licenses available and only selects one, then two other licenses will be available at all times for other projects to use.
Do not run multiple job profiles simultaneously on the same computer: When checked, this will prevent Harvester Server from running multiple, simultaneous collections on the same computer. If unchecked, multiple collection jobs will be run on the same computer if they are queued in that manner.

SUBMISSION OPTIONS

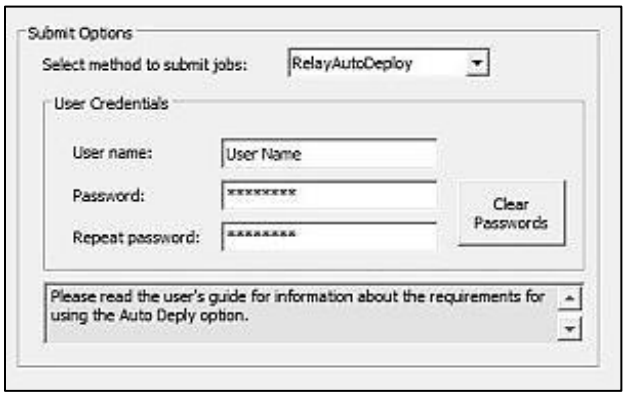
Users can select between using Harvester Server's ***Relay*** or ***PsExec*** to launch collection jobs in a specific project. If ***PsExec*** is selected, the user will see the option to enter credentials for the target system(s).
NOTE: Credentials for target systems can be added later on, but they will have to be entered for each computer individually.

PSEXEC is a command line based administration tool that enables the remote execution of processes on other systems. PsExec comes in the *PsTools* suite commonly used by IT professionals and administrators. Collection jobs submitted in a project can be launched using PsExec, provided the user has credentials for each of the individual target systems.

HRELAY is a program that can be used in place of PsExec to launch collection jobs that are submitted in a project remotely. HRelay has to be installed on computers that are being collected from.

NOTE: Using HRelay requires access to target computers from the shared setting that Harvester Server is located. HRelay can be initiated via Group Policy, or used in conjunction with Relay Auto Deploy.

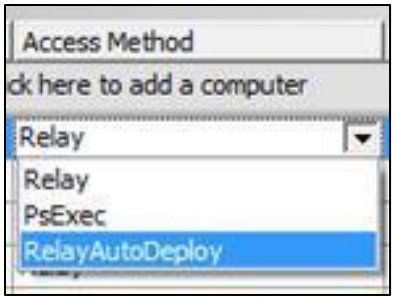
RELAY AUTO DEPLOY option is selected when users want to remotely launch jobs on target computers using Harvester Relay and it isn't already installed and running. Users will be prompted for credentials after ***RelayAutoDeploy*** is selected from the project options dropdown.



When ***RelayAutoDeploy*** is set as the default job **Submit Options** the following steps will be taken for the target computer:

- Harvester Server pings the target computer to check if it is online.
- Harvester Server executes a PsExec command to copy the HRelay file to the target computer \Windows\System32 folder.
- Harvester Sever executes a PsExec command to start HRelay.
- If the above are successful, Harvester Server will send a request to Relay on the target computer to launch the selected Harvester profile.

RelayAutoDeploy can be selected for individual target computers as the *Access Method* in the *Computer List* grid from the dropdown if another option was set by default.



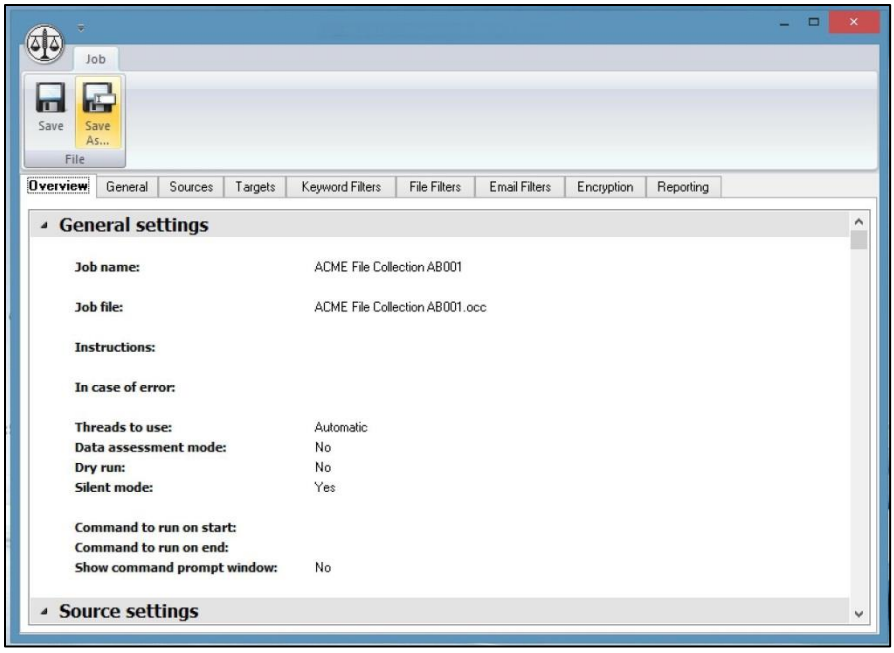
Select **Save** in the project changes portion of Harvester Server and you will have successfully created a project.

Create a Profile

Upon starting Harvester Server, you will first see the main interface screen.
From here, you will be able to create and manage both projects and job profiles.

Opening Harvester Server for the first time, the job profile list will be empty. Click the **New** button 

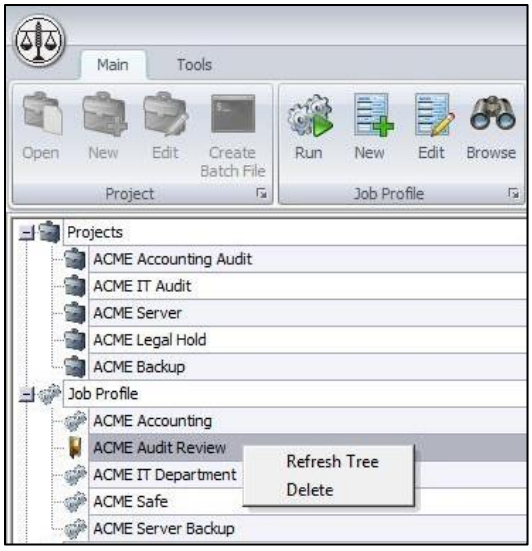
and wait for the job creation screen to appear. Enter the name you would like to assign to the job.
Once a job is defined, you can choose to immediately run the job, or save it for an automated collection or future collection project.



The ***Overview*** tab provides easy access to all the job profile settings for a saved job. This eliminates the need to click each tab as all the settings are summarized in the Overview tab.

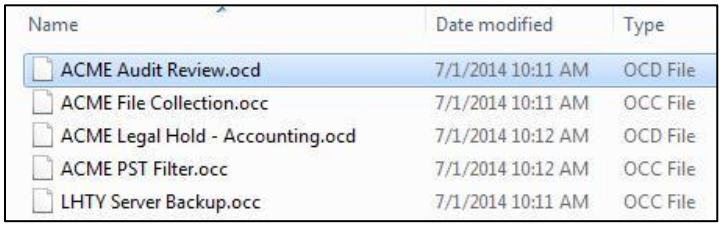
Previously created Job Profiles (both from Harvester Server and Harvester Portable) can be imported to the job profile list in Harvester Server, or they can be dragged and dropped from Windows Explorer into the Job Profile tree area.
Job profiles (.occ) are stored in the _occ directory in the HARVESTER sub-folder of the Harvester Server directory. Once job profiles are copied and pasted into this folder, users will need to refresh the ***Job Profile Tree*** by right clicking and selecting ***Refresh Tree***, to be able to access them directly from Harvester Server.

Users can remove old or unwanted job profiles by right clicking and selecting ***Delete***.



When a job profile is removed, it is not permanently deleted. To recover a removed job profile, follow these steps:

- Browse to the _occ folder located within the Harvester folder of Harvester Server.
- Change the file extension of the removed job or jobs that you would like to recover from ***.ocd*** to ***.occ*** .

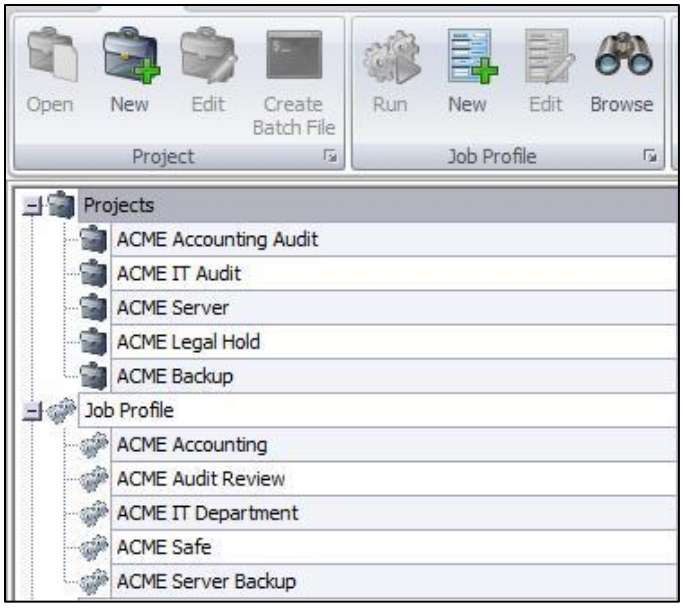


- Right click and select ***Refresh Tree*** to refresh the job profile tree and use the recovered jobs profiles.

EDITING A JOB

After a job profile is created or imported, it can be edited at any time. To edit a job profile, follow these steps:

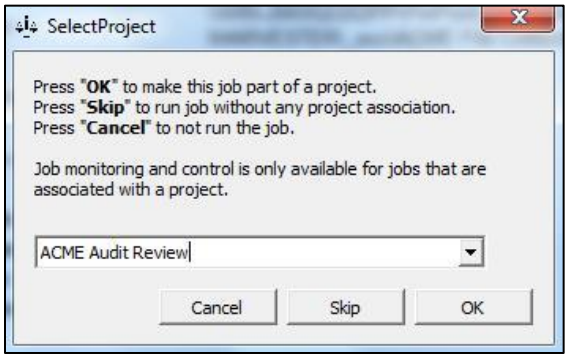
- Select the job profile form the job profiles list and click the ***Job Edit*** button on the toolbar.
- Once the job editing screen appears, make changes to the job profile as needed.
- When finished, click ***Save*** or ***Save as*** to save any changes made to the job profile.
- Exit from the job editing screen, Harvester Server will reappear and any changes made will be effective immediately.



Running a Harvester Server Profile

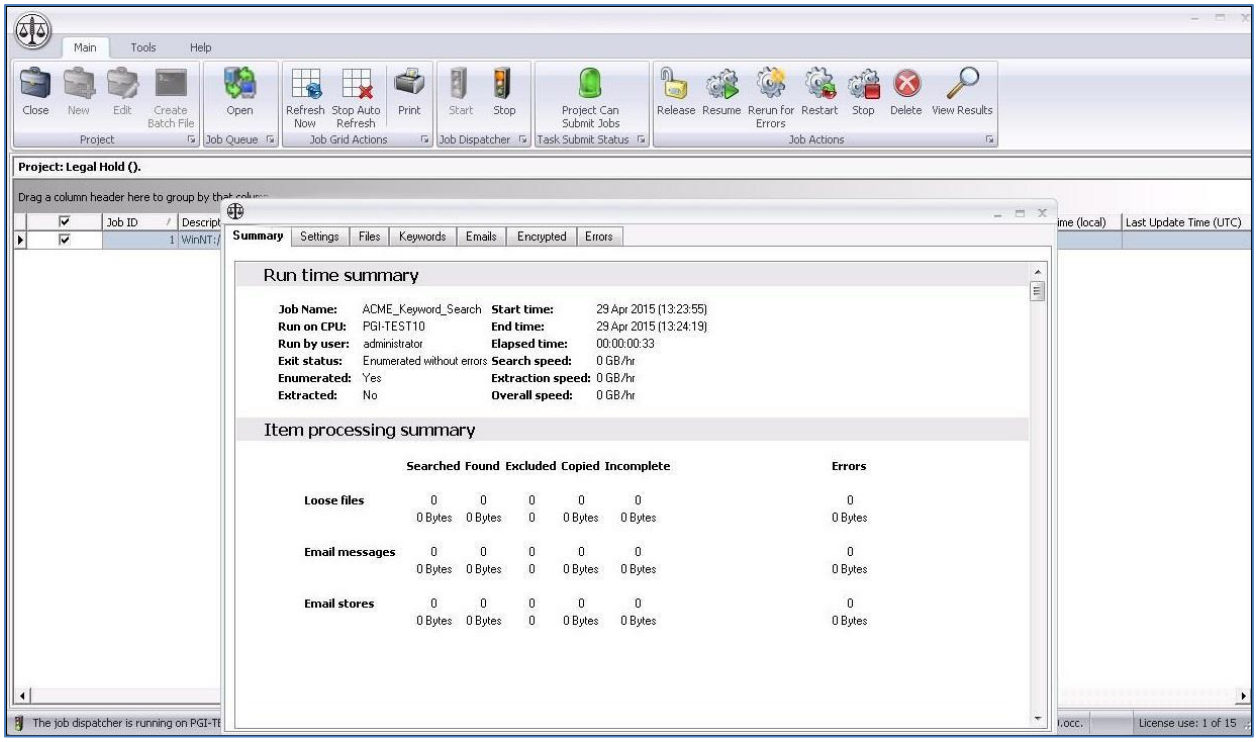
Reference video: ***[Running Harvester Server 5.1 Locally]***

To run a Harvester Server job, click on the profile from the job profiles list and click **Run**. A prompt will appear requiring you to select a project to associate the job with or to skip and let the job run independently.

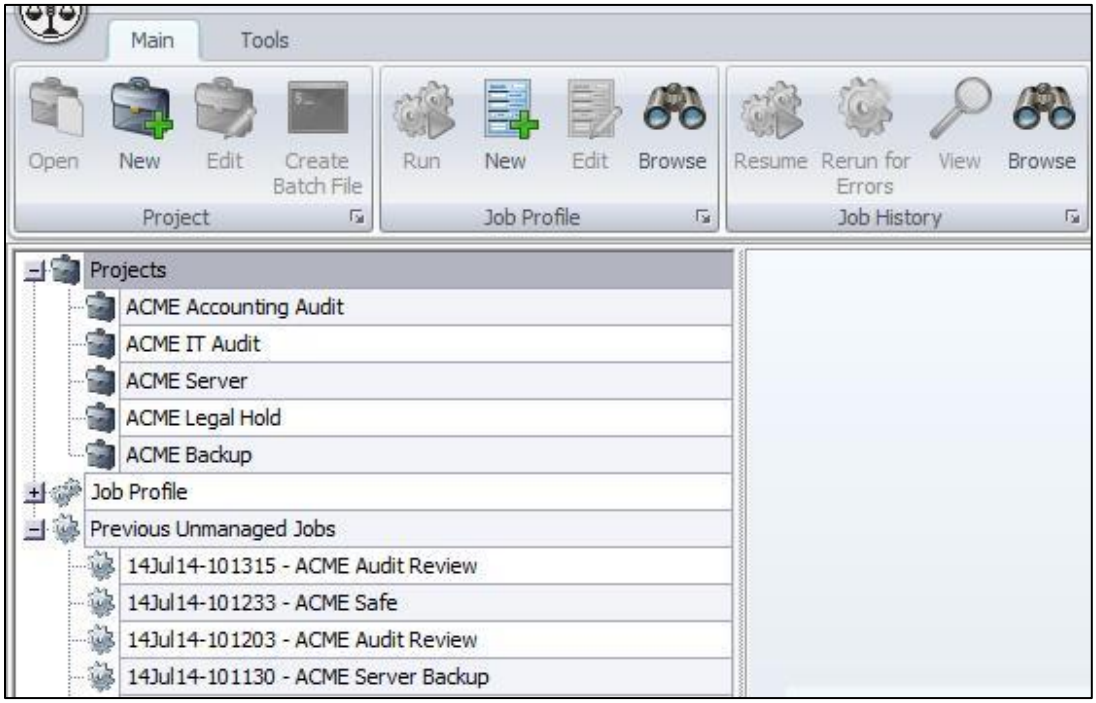


To run a job independent of any projects, select ***Skip***. A window will appear that requires the user to select **OK** to continue. Once the job begins, the Harvester progress console will appear and provide important feedback as well as real-time statistics.

If you would like to associate the job with a project, you must select which project from the dropdown menu of the prompt, and then select **OK**. A window will appear that will require you to click **OK** to confirm that you want to run that job. Once the job begins, a Harvester Server progress console will appear and provide important feedback as well as valuable statistics. You can also open the project that the running job is associated with and receive real-time statistics from within the project manager.

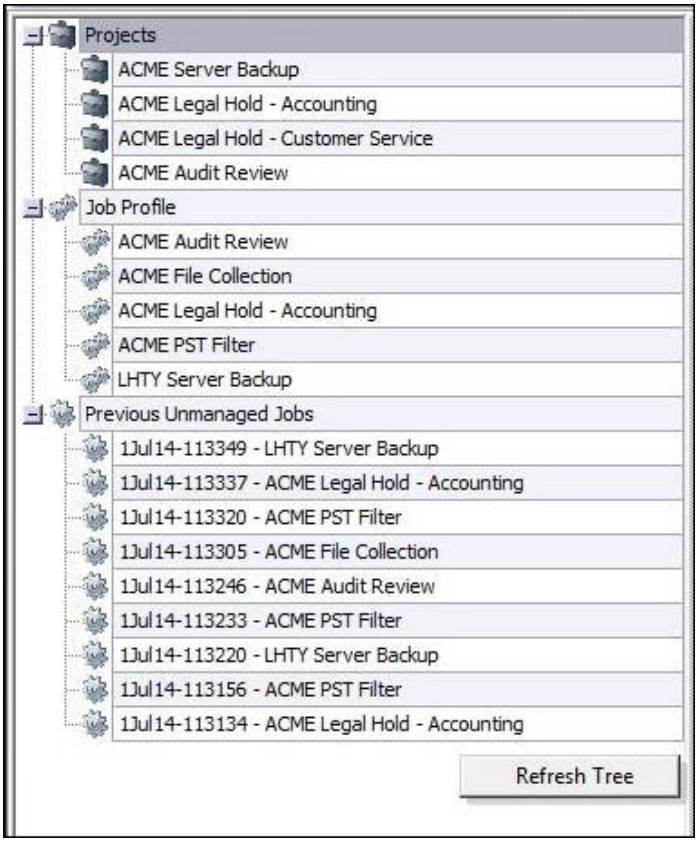


NOTE: If run independent of a project, job results can be viewed by selecting the job from the ***Unmanaged jobs*** list, then selecting ***View Results*** from the toolbar.



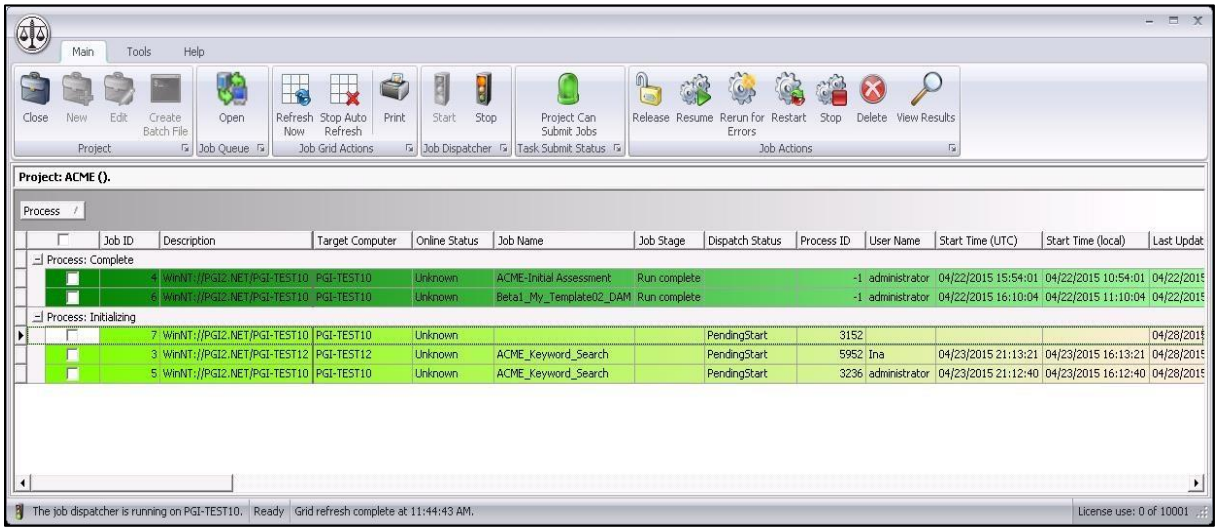
The Previous Unmanaged Jobs tree does not refresh automatically when a job is completed. To view the latest job that has been run, right click in the tree area and select refresh.

To view history statistics for Jobs that have been run outside of Harvester Server or in a previous version, you can browse out to _jobfile.scj files by using the browse button in the **Job History** section of the toolbar. _jobfile.scj files can also be dragged and dropped into the Previous Unmanaged Jobs area and viewed. Job .scj files can be found in the logs folder of each job.

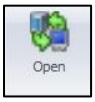


Project Manager - Creating Jobs

Selecting a project in the project list and clicking ***Open*** while having will bring you to the Project Manager screen. This screen contains real time statistics of running and/or previously run jobs.

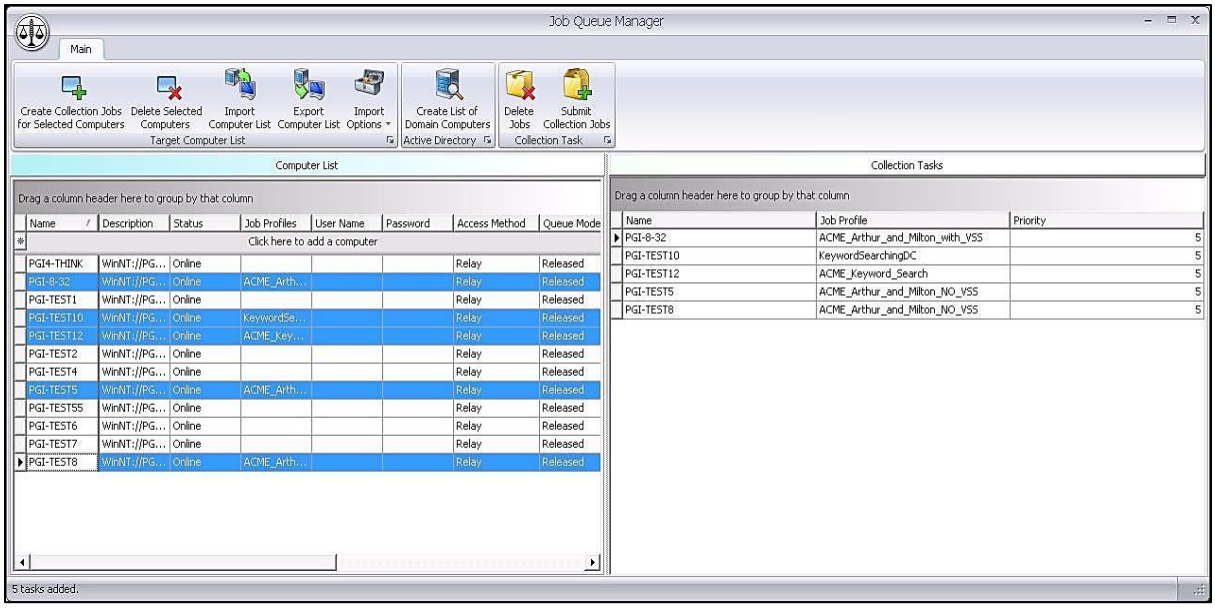


Reference Video: [Harvester Server 5 Advanced Part 2 - Create and Submit Collection Tasks]

When first opening a project, the Project Manager screen will be empty until collection jobs are submitted to it. To start launching jobs on remote systems, click the ***Job Queue Manager*** button: 

JOB QUEUE MANAGER

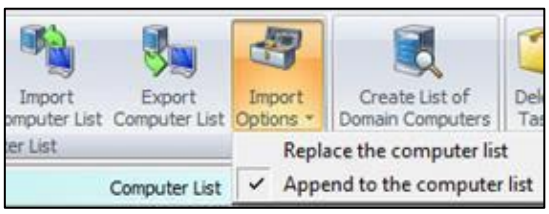
The Job Queue Manager is where the user can view all the computers and jobs available to work with and also set up, review and submit any collection jobs.



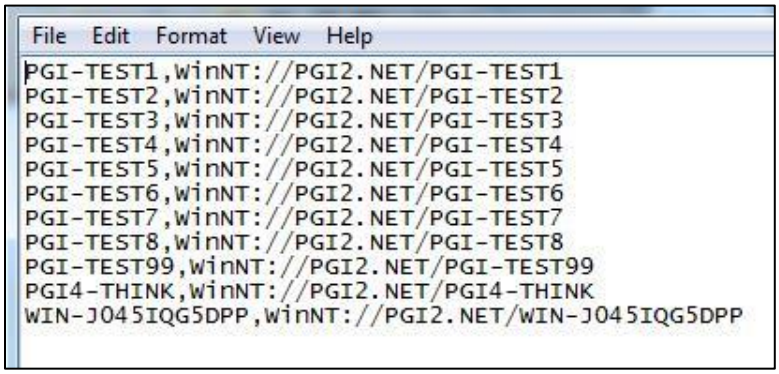
To get started launching jobs on remote systems, you first need to add the target systems to your Computer List Grid. Below are different options for adding target systems to your Computer List Grid.

1. By selecting ***Click here to add computer*** above the Computer List Grid, users can manually add computers to their computer list. Users will have to have computer name and credential information for computers that are added manually.
   NOTE: Harvester Server will not be able to detect whether the information entered is a legitimate system or not, so make sure the system information is correct.
2. Importing a computer list is easy with Harvester Server. Simply click ***Import computer list*** and browse out to your computer list file and click ***Open***. The computers in this list will be added to your Computer List Grid.
   NOTE: Computer list files need to be formatted in comma separated value (**.csv**) files to be successfully imported.
3. Creating a computer list with Active Directory is a good option for those who do not have a previous computer list. Clicking ***Create computer list from domain*** will create a working computer list file of all the computers on the domain and then export that list to the export folder. From there, users can select 

4. ***Import computer list***, browse out to the exported computer list located in the export folder of Harvester Server, and import it into the Computer List Grid.



Creating a computer list- There are two ways to create a computer list without using Active Directory search, using a text document or using Microsoft Excel. To make a computer list using a text document, open a new text document, then type the name and description of the target system separated by a comma as shown below.



- To enter more than one computer, hit the **Enter** key and repeat the process. When finished adding all the target systems, save to a location accessible by Harvester Server.
- To create a computer list using Excel, open a new Excel spreadsheet. Next, type the computer name in the first column and the computer description in the second column. To add more than one computer, hit the **Enter** key and repeat the process. When finished, save the spreadsheet as a .csv file to a location that Harvester Server can access.

CREATING COLLECTION JOBS

A collection task is a job that an assigned target system is to complete. There are many different scenarios where custom collection jobs need to be created.

Reference Video: Harvester Server 5 Advanced Part 2 - Create and Submit Collection Tasks

A common scenario is submitting collection jobs to the target systems that are currently online, and waiting to submit for the offline systems. To create collection jobs for the scenario follow these instructions.

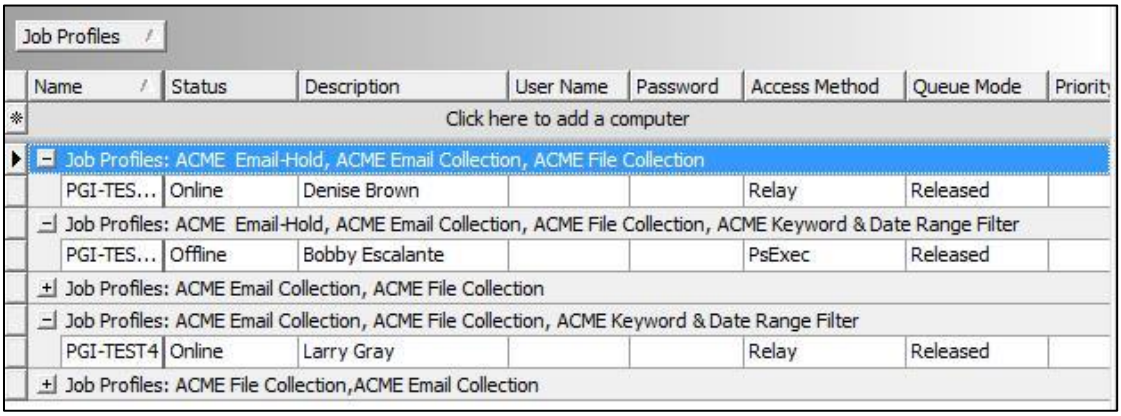
- Select all computers in the Computer List Grid.
- Right click and select **Ping Selected Computers** to check whether the target systems are online.

- Click and drag ***Online status*** to the gray area above the grid to sort target systems by online status.
- Select online target systems that need to be collected from and either click create a collection task in the tool bar, or right click and click ***Create collection task*** from the dropdown menu.
- After reviewing the collection task list, if it is correct, click ***Submit collection jobs***.

Another common collection scenario is having different jobs for the target systems within the same project. To create specific collection jobs in the scenario, follow these instructions

Make sure each target system has the correct box checked.

- Drag the **Job Profiles** field into the gray area above the grid to sort the computer list by the job profiles selected for the target systems.
- Select the group of jobs you would like to run at this time to expand the computer list that fits the specifications.

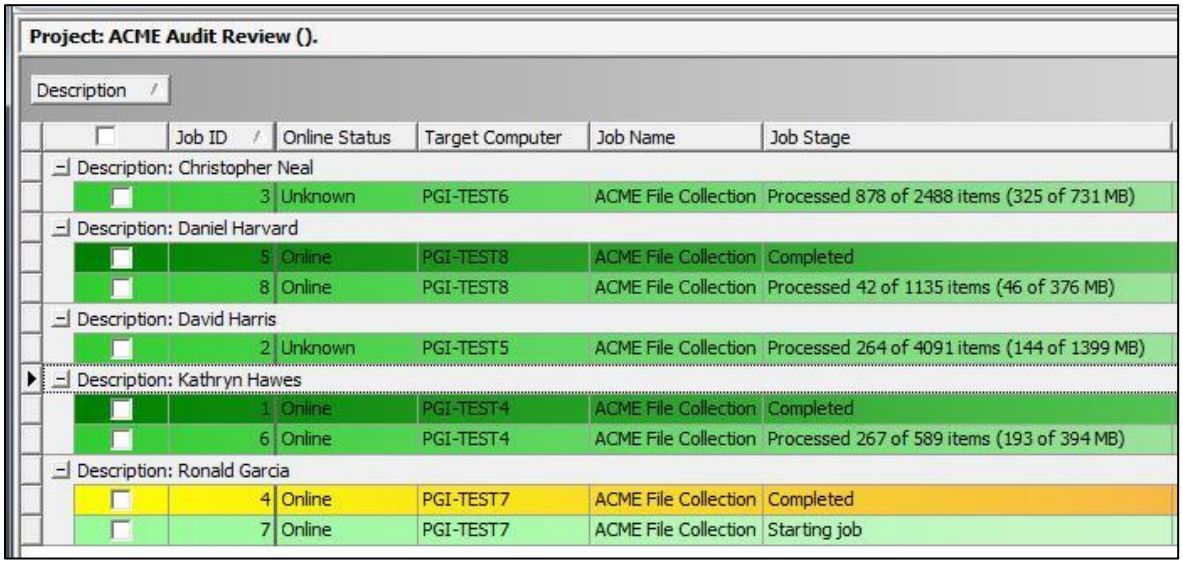


1. Select all computers within the sub category and click **Create collection jobs** in the tool bar or right click and select **Create collection jobs** from the drop down menu.
2. After reviewing the collection jobs list to make sure it is correct, click **Submit collection jobs** from the tool bar.

After submitting collection jobs, the next step is monitoring and managing them in the Project Manager.

Project Manager - Managing Jobs

Once jobs are submitted to Harvester Server, they can be tracked and managed from the Project Manager.



Reference Video: [Harvester Server 5 Advanced Part 3 - Job Manager Features]

When tracking a job in the Project Manager, there are different stages it will go through that provide important information on how the job is progressing. Below are the titles and brief descriptions of the job stages that will appear when a collection task is running or completed without any errors.

- ***Pending Start***: This stage occurs directly after the job is submitted or resumed. During this stage, the collection task is being sent out to the target system and tells Harvester Server where to begin collecting.
- ***Pending Stop***: This stage occurs when a stop command is issued to a collection task.
- ***Start Issued:*** This stage occurs when a start command has been received by Harvester Server but the job is still initializing. Process ID from the target system should be available at this time.
- ***Enumerating:*** This stage occurs when Harvester Server is running through its first stage on the target system. During enumeration, Harvester Server goes through the target sources and marks each file that fits the collection task criteria for processing. In this stage, there will be real-time statistics on the number of files found and excluded for processing.
- ***Processing:*** This stage occurs when Harvester Server is going through its second and final stage on the target system. During processing, all files that were marked for

copying during enumeration are copied to the target location. This stage will provide real-time statistics on the number of files that have been processed.

- ***Completed:*** This stage occurs when Harvester Server has completed the collection task without any errors. To view more in depth information about the completed job, select the job and click ***View results*** from the tool bar or right click on the job and select ***View results*** from the dropdown menu.

Collection jobs will not always run through without any problems. These Harvester Server job stages are meant to help locate the problem and get the user back on track faster. Below are the titles and brief descriptions of the job stages that occur when a problem exists.

- ***Initialization failed:*** This stage occurs when collection task information was communicated to Harvester Server, but Harvester Server failed to start enumerating and processing the data.
- ***Start failed:*** This stage occurs when a stopped or paused job is issued a resume command, but Harvester could not resume the job. This stage can also occur when a job is first started and the collection task is not communicated to Harvester Server.
- ***Stop failed:*** This stage occurs when a stop command was issued but the collection jobs could not be stopped.
- ***Timed out:*** This stage occurs when the collection task in progress has not received update information from Harvester Server for ten minutes. This is usually caused by a problem with the target system such as being disconnected from the server, powering down or restarting. Once timed out, the collection task will move to a stalled job status.
- ***Stalled:*** This stage will occur directly after time out and triggers the Harvester Server *AutoResume mode*, which will attempt to resume the job periodically without any user input. Harvester Server Dispatcher will continue to check every 15 seconds (ping) and see if computers associated with a stalled job are available. If the computer comes back online and the conditions correct a job restart will be issued.
- ***Completed with errors:*** This stage will occur when a collection task has finished, but there were errors. To re-run the collection task just on the files that encountered errors, select the job and click ***Re-run for errors*** from the tool bar, or right click the job and select it from the dropdown menu.

PROJECT MANAGER – TOOLS

Below is a list of each action in the Project Manager and a brief description of what it does.

- ***Close:*** Clicking this will close the current project and return you to the main Harvester Server interface. Jobs will continue to run\launch as long as the project allows it and job dispatcher is running.
- ***Open Job Queue Manager:*** Clicking this will open the Job Queue Manager. From here, users can create collection jobs and submit them to Harvester Server.
- ***Refresh Now:*** Clicking this will force an immediate update of information within the Project Manager grid.
- ***Stop/Start Auto Refresh:*** Clicking this will toggle turning Auto Refresh on and off. When Auto Refresh is on, the information within the Project Manager grid updates every 5 seconds on its own. When Auto Refresh is off, the Project Manager grid will not refresh unless it is turned back on or ***Refresh Now*** is clicked.
- ***Job Dispatcher Start and Stop:*** Job dispatcher is a program that runs in your system task tray that allows Harvester Server to submit jobs when it is both open and closed. Selecting ***Stop*** will close Job Dispatcher and will prevent jobs from being launched completely. Selecting ***Start*** will open job dispatcher and allow jobs to be run.
  NOTE: Switching this setting will affect every project within Harvester Server. Once job dispatcher is started, all projects will be able to run jobs.
- ***Jobs Submit Status:*** Clicking this will allow the user to toggle between allowing the project to submit jobs and not allowing the project to submit jobs. This is useful when stopping job dispatcher is not an option because other projects need to continue to run jobs.
- ***Release Job:*** Clicking ***Release Job*** will allow the selected jobs that are being held to run. This will not affect any jobs that are already released.
- ***Resume Job:*** Clicking this will resume the selected jobs from where they last stopped. This can be used to resume jobs that stopped due to being in data assessment mode. Jobs that have finished successfully cannot be resumed.
- ***Rerun for Errors:*** Clicking this will restart selected jobs from the beginning, retrying all of the errors that occurred during the job.
- ***Restart Job:*** Clicking this will cause the selected jobs to restart from the beginning and they will lose any work that has already been done.
- ***Stop Job:*** Clicking this will stop selected jobs that are running. These jobs are able to be resumed.
- ***Delete:*** Clicking this will delete selected jobs from the Project Manager screen. Collection jobs currently in progress cannot be deleted.
  NOTE: Once deleted, collection task information can only be found in the logs area of

the specified collection task. Users will not be able to view the information from the Project Manager again.

- ***View Results:*** Clicking this will bring up the history information for the selected job.

Harvester Server uses Job Dispatcher to submit and run collection jobs even when the main application is closed.

Job Dispatcher is unique, as it only runs on the system that first opens the instance of Harvester Server. Harvester Server can still be opened and collection jobs launched on other systems. These would be submitted to a job queue and launched from the first system's Dispatcher.

Job Dispatcher start and stop options will also be unavailable to users who start Harvester Server after Job Dispatcher is currently running on another system.

If Job Dispatcher is shutdown on the first system, each system with Harvester Server currently open will receive a warning that Job Dispatcher is not running. Those users will also have the option to start Job Dispatcher on that system.

When Job Dispatcher is started on another system, other Harvester Server applications will lose the ability to start and stop Job Dispatcher again and notification windows will no longer appear.

PROJECT MANAGER GRID

The Project Manager grid shows all the information about collection jobs that are within a project.

The fields within the Project Manager grid can easily be customized to view the most important information available with these methods.

- The entire grid can be reorganized in ascending or descending order by information from within a column by left clicking the title of the column.
- The columns of information can be put into the desired order you want by left clicking, dragging individual columns to the right or left, and then dropping them where you would like them.
- Columns can be reorganized into greater detail by left clicking, dragging the column into the gray space directly above the Project Manager grid and dropping it there.
- By default, some columns will not be able to show all text within the field. To view all the text within a column, simply left click the edge of the column and drag to expand it, or you can double left click the right edge of the column to automatically expand it to fit the largest line of text currently in the column.
- Information from within the grid can be broken down in greater detail by selecting the  in the right corner of an information column and selecting ***Custom.***

Users can also customize the colors that the represent stages and status of jobs by going to the tools section of the Project Manager and selecting the ***Grid colors*** option. From here, the user can either edit each pair of colors individually or complete groups of colors by editing the first in each group and then selecting ***Edit group***.

INFORMATION COLUMNS

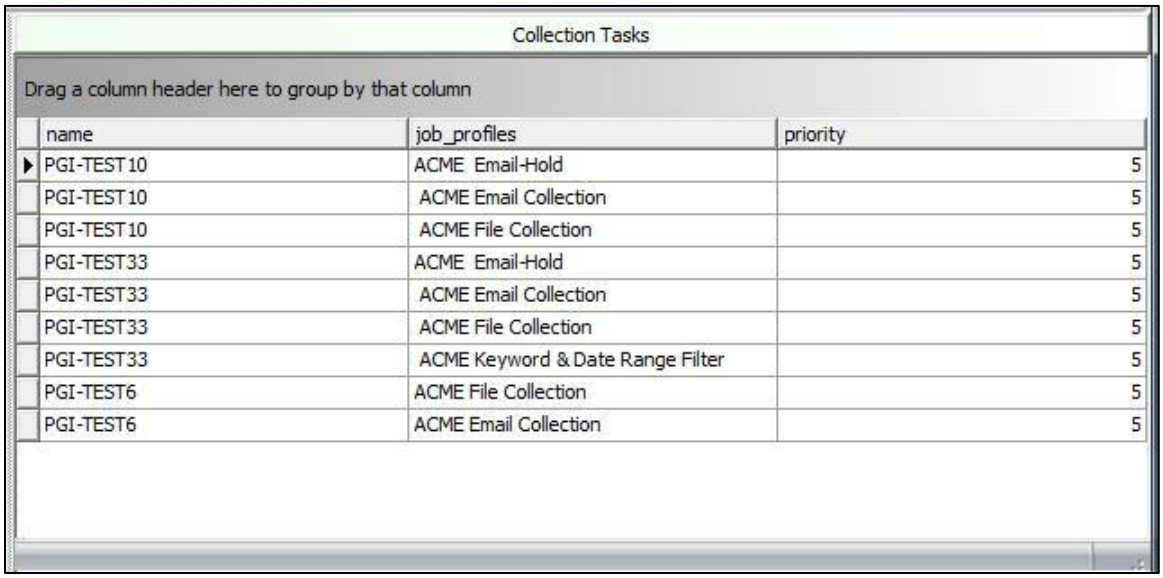
Each column of the Project Manager screen displays different information about the submitted collection jobs. Below is a table containing the name and a brief description of each column within the Project Manager grid.

Column Name	Description
Job ID	Job ID is a number assigned to each job as it is submitted.
Description	Custodian name and other important identifiers.
Target Computer	Computer that is the target of the collection task.
Online Status	Displays whether the target system is online or not.
Job Name-profile	Name of the job that is being monitored.
Job Stage	Completion status of the job being run.
Dispatch Status	Current status of the job dispatcher.
Process ID	Gives the process ID that the target computer assigns to the running job.
Process	Stage of the job being run.
Custodian	This is a legacy field from Net Harvester that is currently not in use
Location	This is a legacy field from Net Harvester that is currently not in use
Username	Username of the target computer.
Computer name	Target Computer
Start time	Time that the job was started.

Last Update Time	Last time that progress for the job was updated.
Elapsed time	Total time the job has been running.
Target Folder	Path to the Target folder of the job.
Logs Folder	Path to the Logs folder of the job.
Loose Files Identified	Number of loose files found.
Loose Files Copied	Number of loose files copied.
Loose Files Skipped	Number of loose files skipped.
Loose File errors	Number of errors found with loose files.
Email Stores Identified	Number of identified email stores.
Email Stores Copied	Number of email stores copied.
Email Stores Skipped	Number of skipped email stores.
Email Store Errors	Number of errors found in email stores.
Messages Identified	Total number of identified messages.
Messages Skipped	Total number of skipped messages.
Message Errors	Total number of errors found with messages.
Dispatch Priority	Priority of the job.
Dispatch release status	Whether the job is held or released.

JOB QUEUE MANAGER TOOLS

Create Collection Jobs: Clicking this button will create collection jobs for any selected computers and the jobs that are checked for them. Once clicked, the collection jobs will be moved over to the collection jobs grid where they can be reviewed before submission.



Delete selected computers: Clicking this will delete the selected computers from the computer grid.

Import Computer List: Clicking this will allow you to select a computer list that will be added to the computer grid.

Export Computer List: Clicking this will export your current computer grid list to a file that can be imported into another project.

Import Options: Clicking this will give you the option of ether appending or replacing the current computer grid list with the imported computer list. If ***Append computer list*** is selected, then the imported computer list will be added onto the bottom of the current computer list. If ***Replace computer*** **list** is selected, then the imported computer list will replace the current computer grid list.

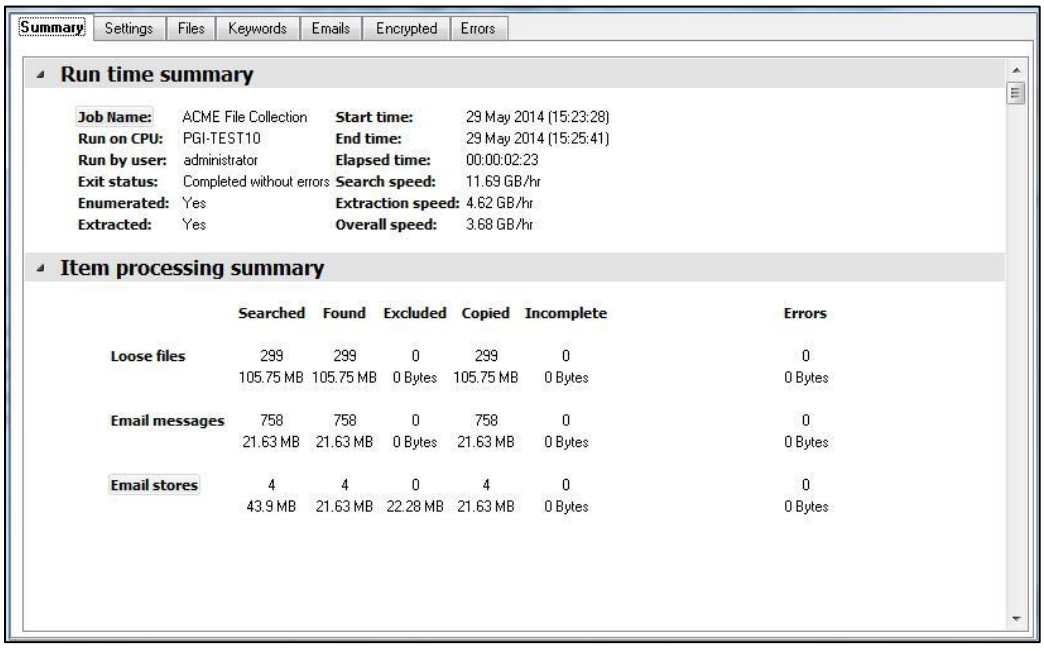
Create Computer List from Domain: Clicking this will make Harvester Server automatically create a computer list of all of the computers in the domain. This computer list can be imported after it is created.

Delete Selected Jobs: Clicking this will delete any selected collection jobs in the collection task grid.

Submit Collection Jobs: Clicking this will submit all collection jobs to be run by Harvester Server. The progress of these jobs can be viewed and managed with live updates from the Project Manager screen.

VIEW RESULTS SCREEN

The view results screen provides more in depth information about a job that has run than the project manager screen provides.



Below is a list and brief description of each of the tabs within the View Results screen.

TAB NAME	DESCRIPTION
Summary	Contains run time statistics and totals for email and loose files categories.
Settings	A snapshot of the job profile settings. This can be very useful if users would like to know if, for example, they chose a setting or included all keywords.
Files	Tally for file types that includes total count and size.
Keywords	Lists total hits for each keyword entry and allows users to launch keyword hits preview.
Emails	Review which mail stores had matching items and the folder location.
Encrypted	Shows list of identified encrypted files organized by type.
Errors	Shows list of identified errors organized by category.

General Tab, Sources Tab and Targets Tab

This table identifies the job profile tabs and how Harvester profile settings are organized:

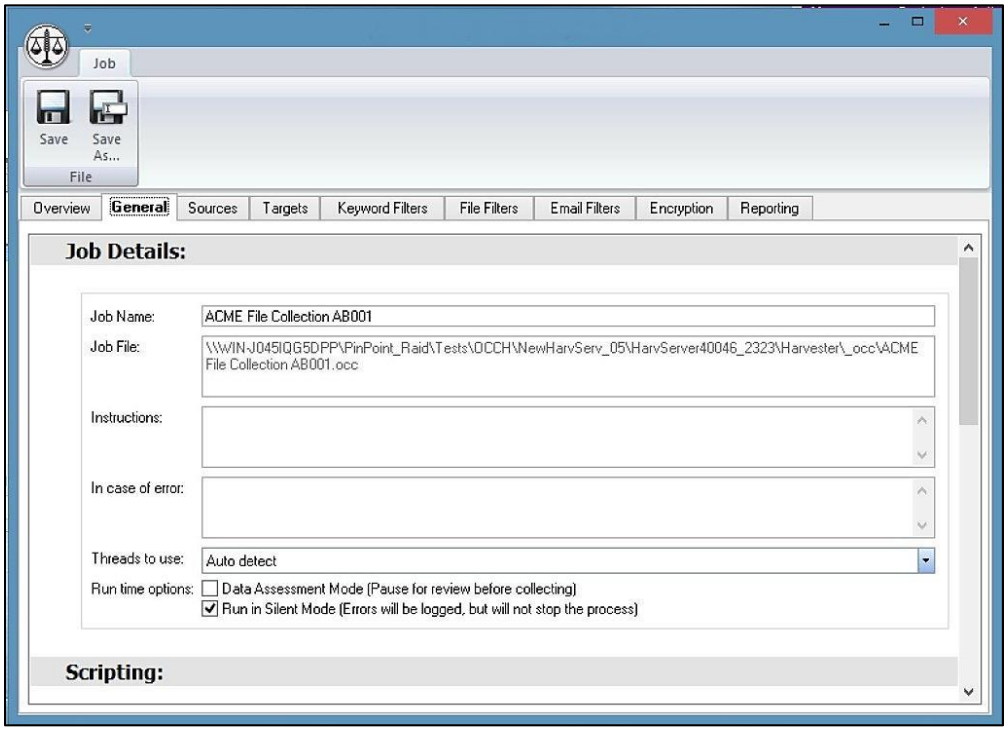
TAB NAME	DESCRIPTION
General	Enter new job profile name and instructions.
Sources	Select sources like local hard drive, file shares, and individual folders.
Targets	Specify where data and logs will be stored.
Key Word Filters	Enter keyword criteria for documents or email.
File Filters	Specify file types and if de-duplication will be used.
Email Filters	Search Microsoft Outlook (PST), Lotus Notes (NSF), and MS Exchange mailboxes.
Encryption	Choose to identify password protected files and how they will be handled.
Reporting	Select from available job logs.

The next sections will identify the individual settings for each tab and a description of their functions.

Reference video: Harvester 5.1 - Creating a Profile OCC

GENERAL TAB

[Harvester 5.1 Advanced Options - Part 1]



Job Name: This is a required field and determines the value used in the [JobName] variable in the file target and job file path. The job name is normally used for the name of a custodian, copy project, or profile (used for multiple systems).

Job File: Current job profile file location. Job profiles are stored in the _occ folder by default; however, users can browse to other job locations by clicking ***Open***.

Instructions: This is an optional description or user instructions that will be displayed in the job list, in a popup window when a job starts, and any time the "i" button is clicked during the run

In case of error: This is an optional field that is displayed when the job starts and again after a job completes if there were errors. It provides contact information for the project manager.

Number of Threads to Use: This option allows you to set a specific number of threads to use for simultaneous copies. If set to **Auto**, the number of threads used will match the number of processors on the machine running the job up to the *MAX_THREADS* value set in the *occ_shell.ini* file in the application directory.

NOTE: You will see diminishing returns when the thread count is set higher than the limits imposed by the system's input and output channels. However, Harvester is designed to limit the impact of these conflicts and to allow for high throughput even on overloaded I/O channels.

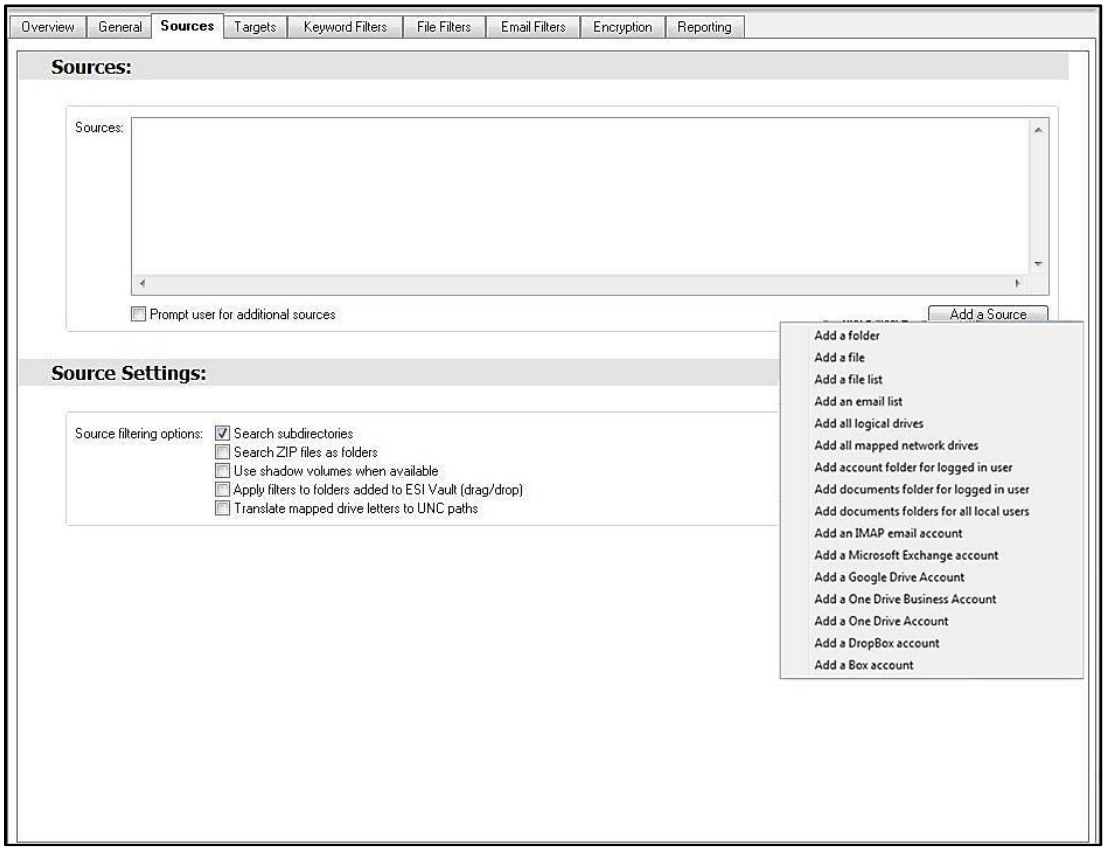
Data Assessment Mode:
Checking this box will stop the job after enumeration so that inventory reports can be generated without the data actually being copied, but leave it in a state where it can be resumed and the files can be copied at a later time.

Run in Silent Mode: Errors that can occur during a project will be logged and this option can often prevent the job from stalling while waiting for a user response (i.e. click the **Ok** button).

Scripting: On occasion, Harvester users would like to launch a job from another application or choose to start a process when a job starts or is finished. This can be accomplished using the scripting options and is covered in detail in this help file.

SOURCES TAB

Data Sources



The data sources window can contain references for drives, directories, individual files or file lists. There are several selection methods available. Click ***Add a Source*** to access the following options:

1. ***Add a folder*** allows users to browse to individual folders.
2. ***Add a file*** is a special file picking window that allows users to select individual files without altering the file time stamps.
3. ***Add a file list*** allows users to select a file list that contains path and filenames or a list of directories. Additional formatting details are listed in the Selecting Data Sources – File List section below.
4. ***Add email list*** allows users to select a file containing entry IDs and the paths to their respective email stores in order to extract individual emails already identified by other software or by a previous run of Harvester.
5. ***Add all logical drives*** inserts the [LDrive] variable that will result in Harvester searching all local logical drives (i.e. C:, D:,E:,)
6. ***Add all mapped network drives*** inserts [MDrive] variable that will result in Harvester searching all locally mapped network locations.
7. ***Add account folder for logged in user*** inserts the [UserAccount] variable that will result in Harvester locating and searching the user account folder for the logged in user.
8. ***Add documents folder for logged in user*** inserts [UserFolder] variable that will result in Harvester locating and searching the documents folder for the logged in user.
9. ***Add documents folders for all local users*** inserts [UserFolders] variable that will result in Harvester locating and searching the ***documents folders of all user accounts on the system.***
10. ***Add IMAP Account*** Inserts ***[IMAP=]*** variable and prompts user to enter criteria for an individual IMAP account. User will need to provide ***1)*** IMAP Server Name ***2)*** Email Account ***3)*** Password ***4)*** Port and ***5)*** Encrypted connection setting.
11. ***Add Microsoft Exchange Account*** Inserts ***[EXCH=]*** variable and prompts user to enter criteria for the web ***1)*** URL ***2)*** User Name, and ***3)*** Password. This option will directly connect to an individual Microsoft Exchange account. It differs from 'Search connected Exchange mailbox' in Email Filter options which uses Microsoft Outlook via MAPI connection
12. ***Users can drag and drop*** files, folders, or drive letters into the Sources field from Windows Explorer. Drive letters and individual emails can also be dragged and dropped to the Sources field from Outlook.
13. Selecting the checkbox ***Prompt user for additional data sources*** will result in Harvester displaying the ***ESI "Easy" Vault*** window. This is commonly used when distributing self-collection kits or jobs from a legal hold notice so custodians can select sources.
14. ***Add Google Drive Account*** inserts GOOGLEDRIVE into the Sources box. This option will connect to an individual Google Drive account. Account login prompts will appear once the job has been started.

15. ***Add One Drive Business Account*** inserts ONEDRIVEBUSINESS into the Sources box. This option will connect to an individual One Drive Business account. Account login prompts will appear once the job has been started.
16. ***Add One Drive Account*** inserts ONEDRIVE into the Sources box. This option will connect to an individual One Drive account. Account login prompts will appear once the job has been started.
17. ***Add Dropbox Account*** inserts DROPBOX into the Sources box. This option will connect to an individual Dropbox account. Account login prompts will appear once the job has been started.
18. ***Add Box Account*** inserts BOX into the Sources box. This option will connect to an individual Box account. Account login prompts will appear once the job has been started.

NOTE: Only a single Cloud source (#14-18 document repositories listed above) may be added per job. *Multiple additions of Cloud sources will be ignored.*

Files and folders below NTFS reparse points such as junction points, symbolic links, and mount points are not accessed or collected by Harvester. If Harvester encounters a folder with the reparse attribute, it will place an entry in a log in the logs folder (*_mountpointss.log*, *_symlinkss.log*). Reparse points can point to a non-existent source because the operating system does not check to see if the source exists. Harvester does not treat symbolic links as folders or files due to these possibilities.

- A mounted drive can contain a symbolic link to a path that also exists on the examiner's machine, leading to the copying of irrelevant data
- A symbolic link can contain a reference to a folder higher in its own folder hierarchy, causing an infinite loop.

NOTE: Files and folders below NTFS reparse points may be accessed and collected by Harvester by changing settings in the **occ_shell.ini**. Located in the **bin** folder, the Harvester **occ_shell.ini** can be accessed with a text editor (such as Notepad), and changed. Changing the **FOLLOW_SYM_LINKS** field from **0** to **1** and saving the document will allow Harvester to follow symbolic links, mount points, and junction points.

Selecting Data Sources - File List

File lists generated from full text search engines, litigation support databases and computer forensic software can easily be imported using the Add a file List option. When relevant files or directories are identified, the file list option provides an alternative to manually selecting or dragging and dropping directories into the data sources field.

By selecting the ***Add a file list*** option, you will be able to browse to the location and select the list to be used. A file list can be any text file (.txt, .csv, .log), so long as the full file path or folder path is the first field in a tab-delimited text file. The list file can contain one file path or directory per line.

Individual files

c:\Documents and Settings\Jon\Desktop\PGP Source\HD\Docs\Articles - Forensic\DCFL Request Letter Format (12 Jun 00).pdf

c:\Documents and Settings\Jon\Desktop\PGP Source\HD\Docs\Articles - Forensic\Digital Evidence Standards (Public).pdf

c:\Documents and Settings\Jon\Desktop\PGP Source\HD\Docs\Articles - Forensic\Digital forensics of the physical memory.doc

Individual directories

c:\Documents and Settings\Jon\Desktop\PGP_Source\HD\Docs\Articles\

c:\Documents and Settings\Jon\Desktop\PGP_Source\HD\Docs\Articles\

\\HSGT-01\shares\sales

\\HSGT-01\shares\marketing

Using _errors log as file list

If errors are encountered, they are written to a file called ***_errors.log***. This log can be used as a file list, which will allow you to reprocess files that resulted in errors during a run. This option is especially useful when files are in use and can't be copied. A common use would be to use the error log to copy open files once they are closed. To process an error log, select the Add a file List option, browse to its location and select ***Open***. If you use the same target folder as the original run, Harvester will reattempt to copy any files that could not be copied previously.

Variables

[LDrive]

This variable (with the brackets) indicates that the program should search all logical drives that are connected to the computer. This includes flash drives, CDs, internal hard drives and RAID devices. It does not include network shares or the device that the Harvester software is running from or copying to.

[MDrive]

This variable (with the brackets) indicates that the program should search all mapped network drives. This includes all drive letters that are mapped to a network location (ex: P:\ (\\netshare\files\johndoe1)). It does not include the drive that the Harvester software is running from or copying to.

[PROMPT]

This variable (with the brackets) indicates that the program, when run, should prompt the user to drag and drop source files, folders and emails into the ESI Vault window. Users can also select ***Prompt user for additional sources***, which eliminates the need for the [PROMPT] variable. Using [PROMPT] as a source allows you to define specific sources, as well as requires the user to specify additional sources at run time. The ESI vault window no longer pops up if no sources are specified

[UserFolder] and [UserFolders]

The **[UserFolder]** variable can be used as a source to add the logged in user's My Documents folder as a source. **[UserFolders]** can be used as a source to add all accessible My Documents folders.

File lists generated from full text search engines, litigation support databases and computer forensic software can easily be imported using the **Add List** option. When relevant file directories are identified, the file list option provides an alternative to manually selecting or dragging and dropping directories into data sources.

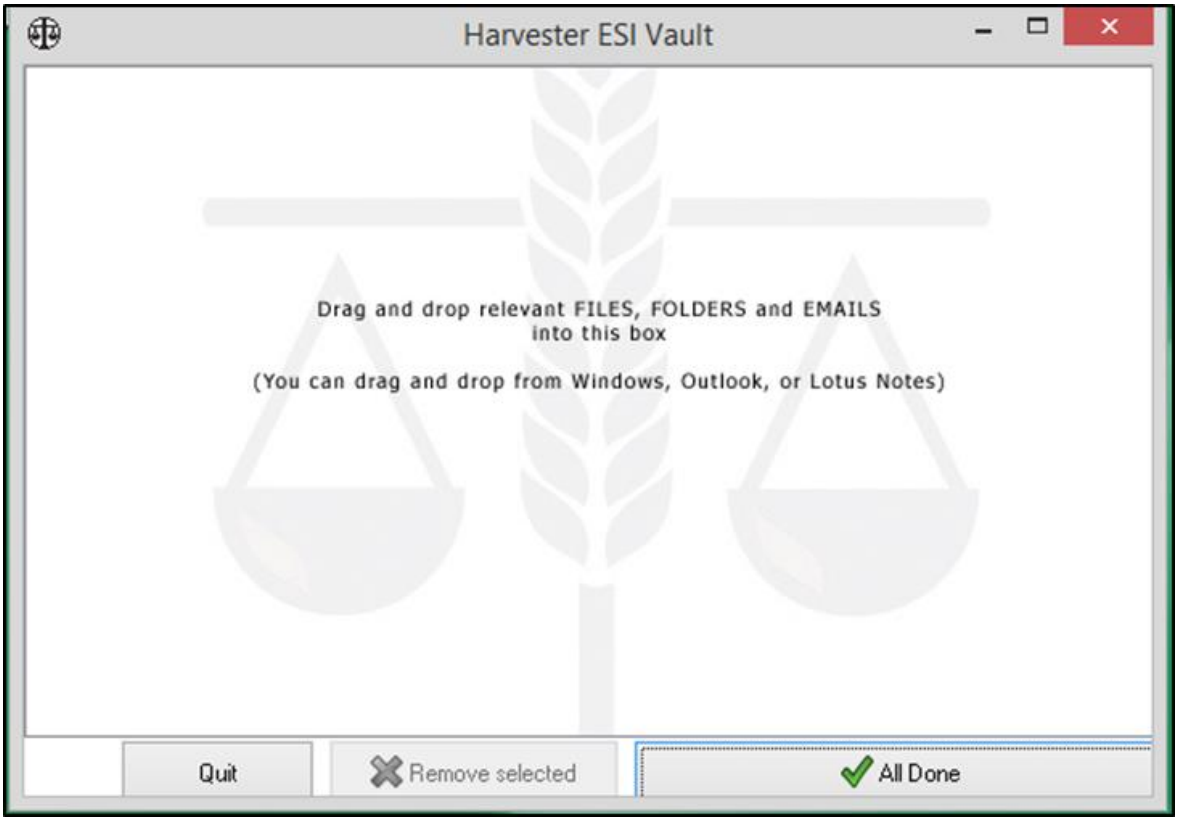
Drag and Drop Source

Files, folders, and Outlook emails can be dragged and dropped into the **Source** window, similarly to the **ESI "*Easy Vault*" Vault** detailed below.

ESI Easy Vault

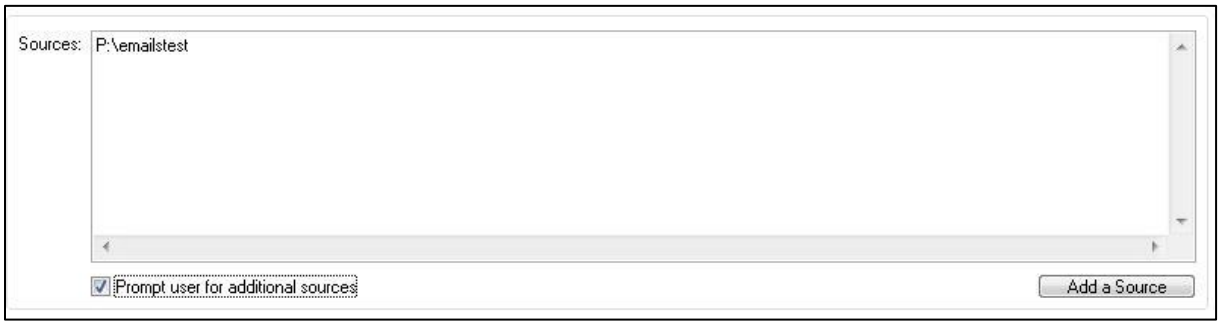
As stated in the last section, the Harvester ESI Vault is commonly used when distributing self-collection kits or jobs launched from Harvester Server for legal hold notice. By providing this simple interface and instructions specific to each job, custodians can easily identify items relevant to a matter. The ESI Vault interface is a window that supports dragging and dropping of the following types of items.

- Files
- Folders
- Emails (must be dragged and dropped from Microsoft Outlook or Lotus Notes)

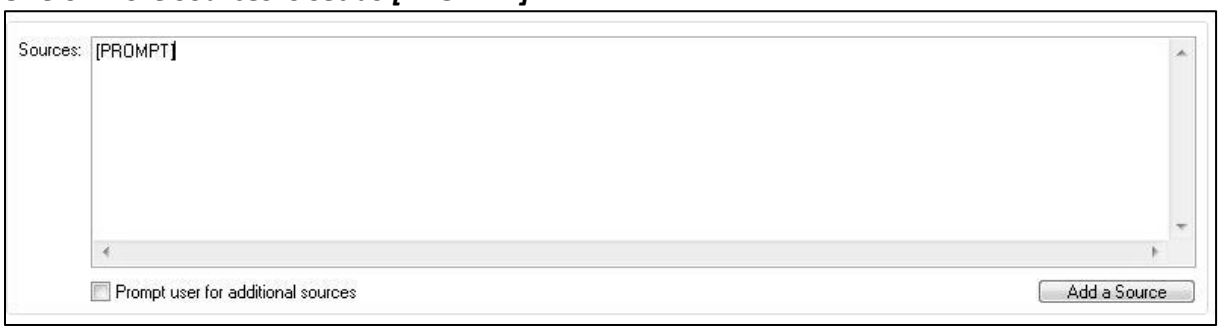


The ESI Vault can be used with local files and folders or network file shares. There are two scenarios that will launch the ESI Vault during a job:

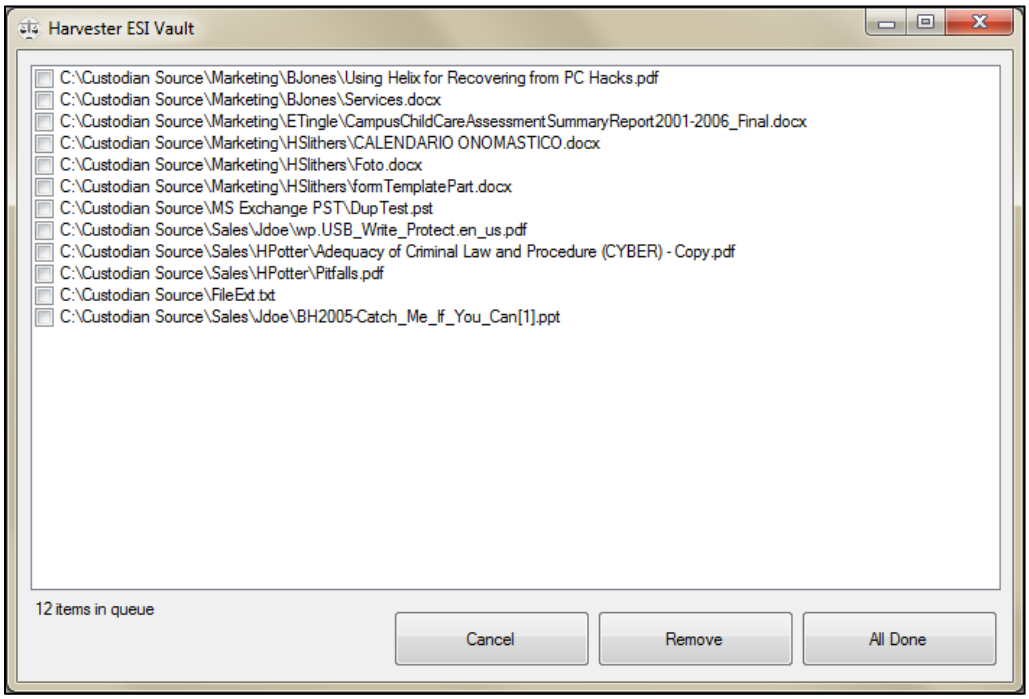
Select 'Prompt user for additional sources' under 'Sources' tab



One or more sources is set as [PROMPT]

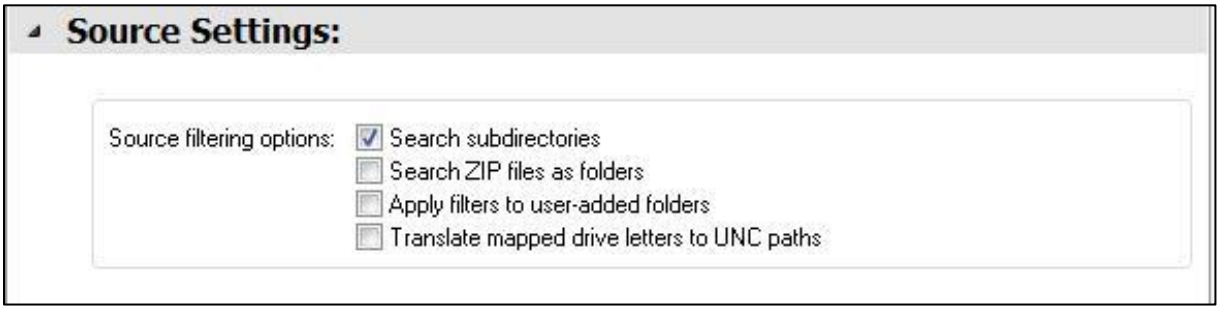


Instruction defined in the Instructions field when creating jobs can be viewed by clicking the Information button.



Any sources that the user adds can be removed by selecting the items and clicking the ***Remove Selected*** button. Pressing ***Quit*** will exit the collection job. Pressing ***All Done*** will add these sources to the job and process them along with any sources added in the *Sources* field at design time.

SOURCE SETTINGS



Search Subdirectories: This option specifies whether subdirectories are searched. Deselecting this option will cause the program to only search for files in the root of the selected directories and ignore any subfolders it encounters.

Search Zip files as directories: This option filters file type, date, and extension, file name inclusion and file name exclusion filters within zip files.

Use shadow volumes when available: Checking this box will cause Harvester to attempt to create a shadow volume of each of your unique source volumes so that files that are in use can still be copied. Harvester needs to be run as Administrator on a Windows Vista or higher computer for this to succeed.

- ***Shadow Copy*** – also known as Volume Shadow Service or VSS – is a technology included in Microsoft Windows that allows taking manual or automatic backup copies or snapshots of computer files or volumes, even when they are in use. It is implemented as a Windows service called the Volume Shadow Copy Service.
- ***Shadow Copy technology*** requires Windows Vista or higher. It also requires the file system to be NTFS in order to create and store shadow copies. Shadow Copies can be created on local and external, or removable volumes by any Windows component that uses this technology, such as when creating a scheduled Windows Backup or automatic System Restore point.

Apply filters to user-added folders: This option specifies whether filters should be applied to folders added via ***Drag-n-drop*** to the ESI Vault by the user. This does not apply to individual files added to the vault.

Translate mapped drive letters to UNC paths: This option may be selected to translate source paths that are on mapped network drive paths to their UNC paths. The UNC path and file name will appear in the *filelist.txt* and *folderlist.txt* in the log folder. This is useful for providing unambiguous source locations for files residing on the network.

IMAP/EXCHANGE SOURCES

IMAP connection allows users to enumerate and collect directly from web based email accounts such as *Gmail*, *Outlook.com*, and *Yahoo*. Harvester searches these email accounts just as it would a local email store. For IMAP filters, see EMAIL FILTERS.

EXCHANGE connection allows users to enumerate and collect directly from an (online) Exchange server using web services. Harvester searches these accounts just as it would a connected Exchange account. For EXCH filters, see EMAIL FILTERS below. The difference between EXCH and Connected Exchange is EXCH is using web services to collect from the server directly, where Connected Exchange is using the connected Outlook Exchange account.

In Harvester's PPLM folder (bin\PPLM) there is a file named "imap_preset.txt", this file contains presets for common IMAP accounts such as Gmail, Yahoo, and Hotmail (Outlook.com). This presets file can be changed to match what the user commonly encounters and saved. When selecting IMAP from the sources button, the presets can be selected from the topmost drop down option.

NOTE: IMAP and EXCH enumeration and collection speeds may vary based on:

- Network speed and connectivity
- Internet speed and connectivity
- IMAP or Exchange server speeds, connectivity, and availability
- IMAP or Exchange server connection, access, and credential settings.

NOTE: For IMAP connections, IMAP needs to be turned on, and in some cases less secure application allowance will need to be turned on. For *EXCH* connections, Web Services will need to be turned on and the Web Services URL will be needed.

CLOUD SOURCES

Cloud source connections allow users to enumerate and collect loose files directly from web based cloud storage accounts such as Google Drive, OneDrive, OneDrive for Business, Box, and Dropbox.

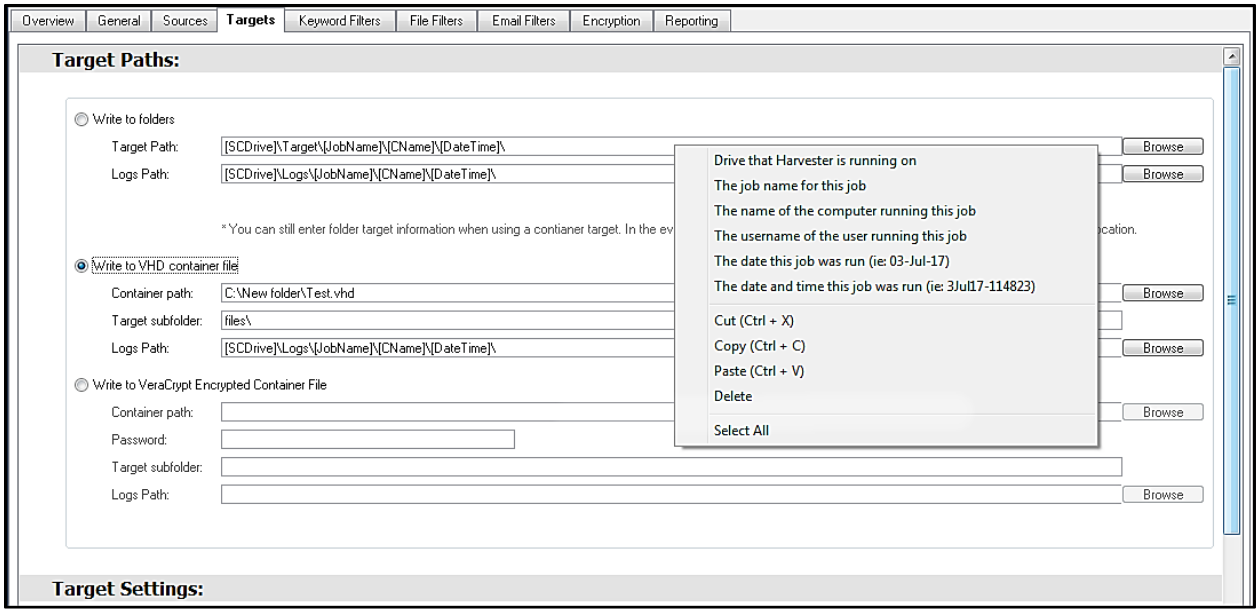
TARGETS TAB

Harvester 5.1 Advanced Options - Part 2

The target path is used to specify where data matching your criteria will be copied. In addition to browsing to an external hard drive, host computer drive or network file share, there are several variables that can be incorporated into the paths. You can also drag and drop a folder location into the field to set the path location.

The variables listed in the table below can be manually entered or you can right-click on the field in the target path to display a drop down of the user friendly descriptions and have them automatically inserted as shown below:

VARIABLE NAME	DESCRIPTION
[SCDrive]	The drive letter that Harvester is running from. (ie: D:)
[JobName]	The name of this job.
[CName]	The name of the computer running this job.
[UName]	The username of the logged-in user running this job.
[Date]	The date the job was run.
[DateTime]	The date and time (to the second) that the job was run.



WRITE TO VHD CONTAINER FILE:

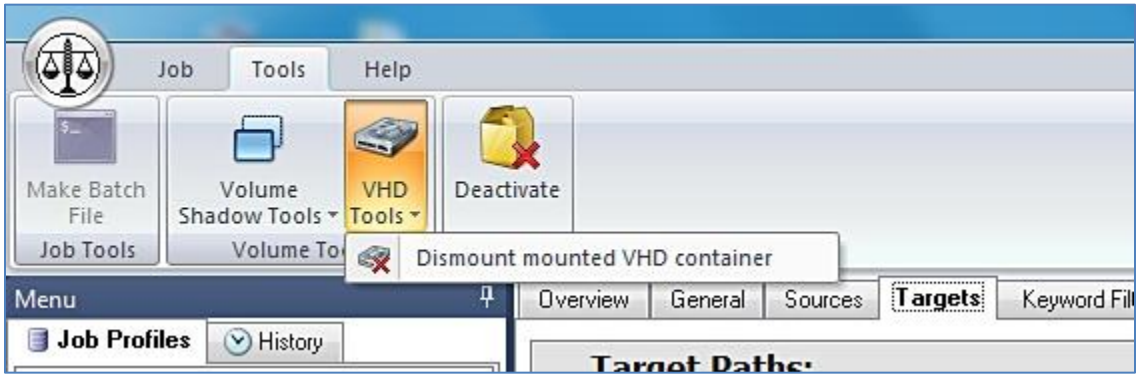
VHD, or Virtual Hard Drive, creates file containers for collected data, keeping all collected data in a single container file for easier transport. A VHD container file acts like any other kind of file, with the exception that it can also act as a hard drive in Windows. Files that have been copied to this virtual hard drive will stay inside the VHD file.
VHD creation is *automatic* when ***Write to VHD file container*** is used as a target.

NOTE: VHD is only supported in Windows Vista and higher. In Windows 8.1, Harvester must be run as Administrator in order to use VHD.
When choosing to copy files to a VHD container, Harvester creates a VHD file at the location specified and formats it like a hard drive. As the Harvester job enters the copy phase, the files are written to the VHD container instead of a target folder. After the job has run, the VHD file itself can later be mounted as a hard drive, either by Harvester or by Windows. The VHD Tools can be found under the Tools tab in the upper ribbon, in the Volume Tools section.

The maximum capacity of the VHD containers used by Harvester is 2 terabytes (2000 GB).

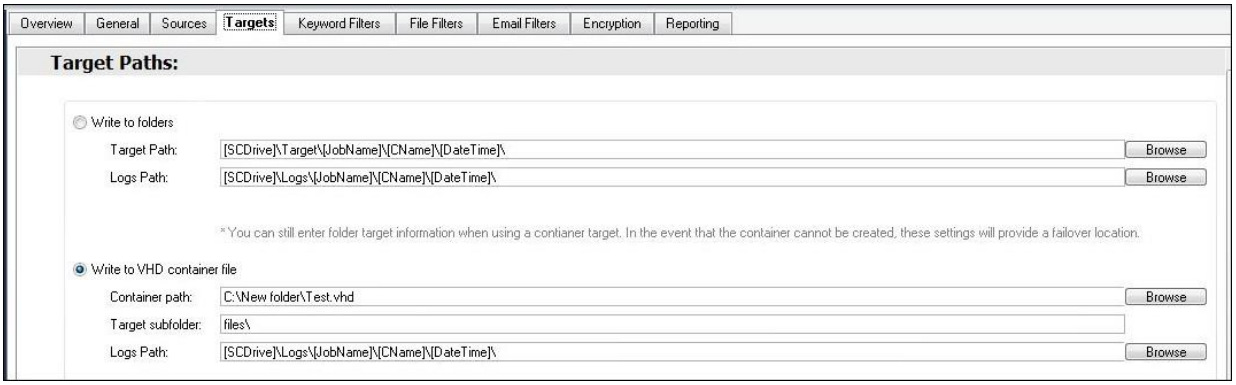
VHD Containers can be mounted to a drive letter by going to the Tools menu and selecting *Mount a VHD container to drive letter*, and mounted to a folder by selecting *Mount a VHD container to a folder*.

After a VHD is mounted, the option to dismount a VHD container is available. VHD containers are automatically dismounted and detached at the end of the job and/or when Harvester is closed.



Under Target Paths,

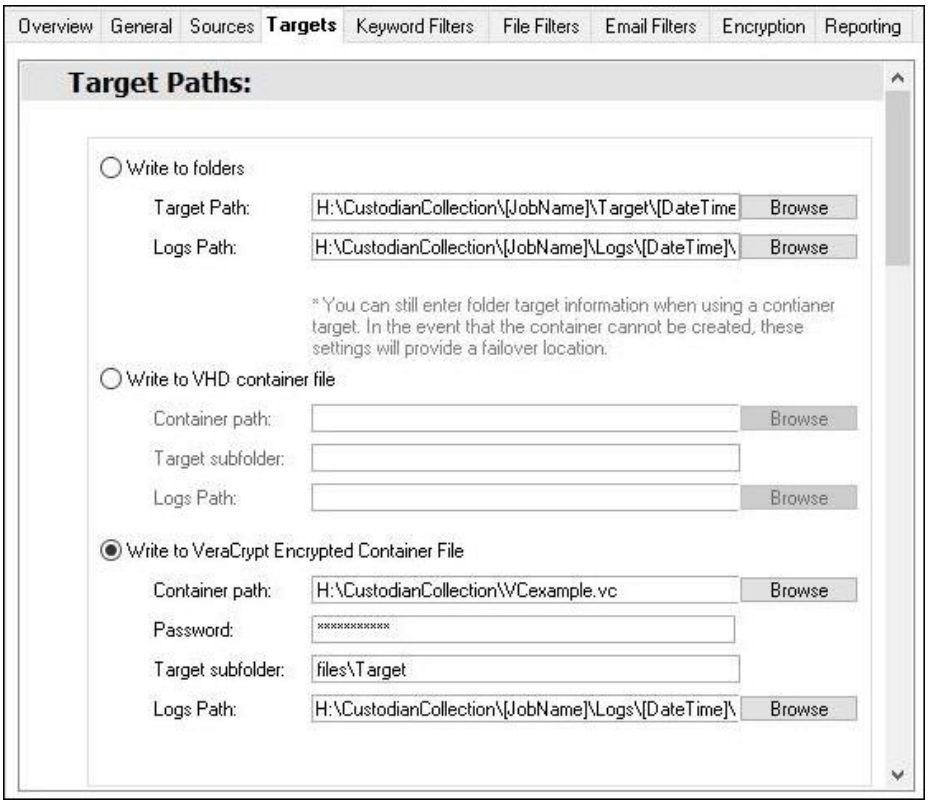
- ***Container path:*** This is the path to the VHD container file.
- ***Target subfolder:*** This is the path to the target subfolder within the VHD container file.
- ***Logs Path***: This is the path to the logs folder, which can be placed in any preferred location.



NOTE: It is recommended that you provide Target and Logs paths in the ***Write to folders*** locations as well; these paths are used as a failover in case a VHD cannot be created at run time.

WRITE TO VERACRYPT CONTAINER FILE:

VeraCrypt creates encrypted file containers for collected data, keeping all collected data in a single encrypted container file for easier transport. A VeraCrypt encrypted container file acts like any other kind of file, with the exception that it can also act as a hard drive in Windows. Files that have been copied to this encrypted container file will stay inside the container file.



VeraCrypt container creation is automatic when *Write to VeraCrypt encrypted container file* is used as a target.

NOTE: Harvester must be run as Administrator in order to use VeraCrypt.

When choosing to copy files to a VeraCrypt encrypted container, Harvester creates a VeraCrypt file at the location specified, and formats it like a hard drive. As the Harvester job enters the copy phase, the files are written to the VeraCrypt container instead of a target folder. After the job has run, the VeraCrypt container file itself can later be mounted as a hard drive, either by Harvester or by Windows with VeraCrypt. The VeraCrypt Tools can be found under the Tools tab in the upper ribbon, in the Volume Tools section.

The maximum capacity of the VeraCrypt containers used by Harvester is 2 terabytes (2000 GB).

VeraCrypt Containers can be created, mounted to a drive letter or dismounted by going to the Tools menu and selecting *Mount VeraCrypt container.*

Logs Path:

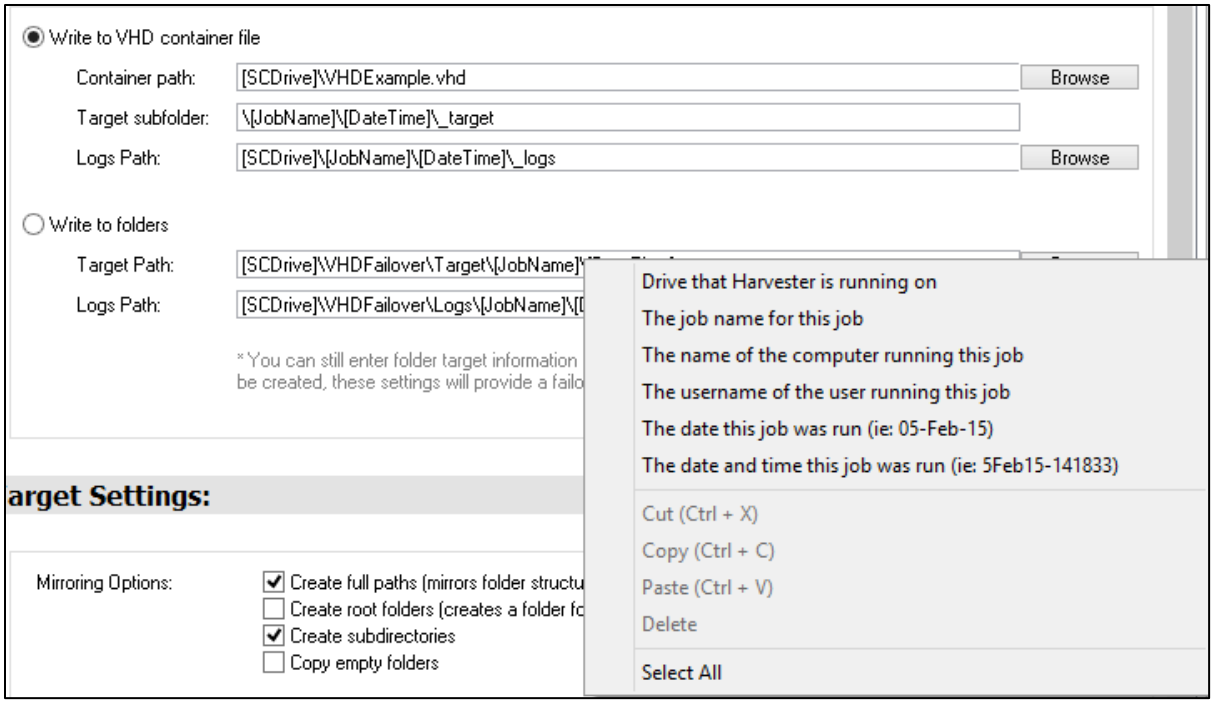
Several logs are created during a collection project and the Logs Path will set where these files are stored. In addition to browsing to a specific folder on a local drive or network file share, the above variables may also be used in the same manner as the Target Path.

NOTE: It is recommended to store logs in a separate path from the Target.
A different Logs folder must be created for each new job to prevent appending data from different jobs, which will lead to serious issues.

Target Path:

The target path is used to specify where data matching your criteria will be copied. In addition to browsing to an external hard drive, host computer drive or network file share, there are also several variables that can be incorporated into the paths. You can also drag and drop a folder location from Windows Explorer into the field to set the path location.

The variables listed in the table below can be manually entered, or right-click on the in the target path field to display a drop down of the descriptions and have them automatically inserted as shown below.



VARIABLE NAME	DESCRIPTION
[SCDrive]	The drive letter that Harvester is running from. (ie: D:)
[JobName]	The name of this job.
[CName]	The name of the computer running this job.
[UName]	The username of the logged-in user running this job.
[Date]	The date the job was run.
[DateTime]	The date and time (to the second) that the job was run.

Logs Path:

Several logs are created during a collection project and the Logs Path will set where these files are stored. In addition to browsing to a specific folder on a local drive or network file share, the above variables may also be used in the same manner as the Target Path.

NOTE: It is recommended to store logs in a separate path from the Target.
A different Logs folder must be created for each new job to prevent appending data from different jobs, which will lead to serious issues.

Target and Logs Path Auto-Check:

Harvester has an Auto-Check feature that will cause the text of the *Target Path* and *Logs Path* to display in red if the respective paths will not translate to actual paths.

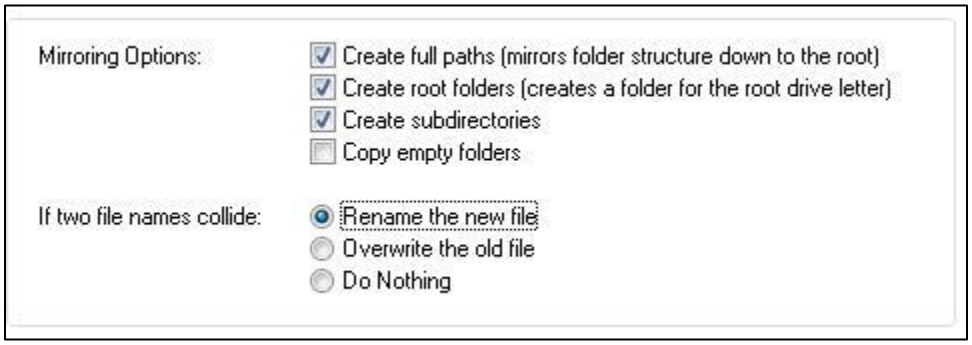
Mirroring Options:

Create Full Paths:
This allows the destination directory to contain a full path of all files and directories that are collected or copied. When selected, the option to *Create Root Folders* is also available.

Create Root Folders:
Checking this option will create a directory for the drive letter or UNC name of the source path. This is useful when the source consists of multiple drives or UNC paths, where each will have a folder containing the files and folders contained therein.

Create Subfolders:
This option is selected by default and matches the directories of the source files. If you would like to copy all source files into a single target folder with no subfolders, then you can deselect this option.

Copy Empty Folders:
This option specifies whether a folder will be created in the target when the source directory is empty or contained no matching documents.



<u>If File Names collide:</u>

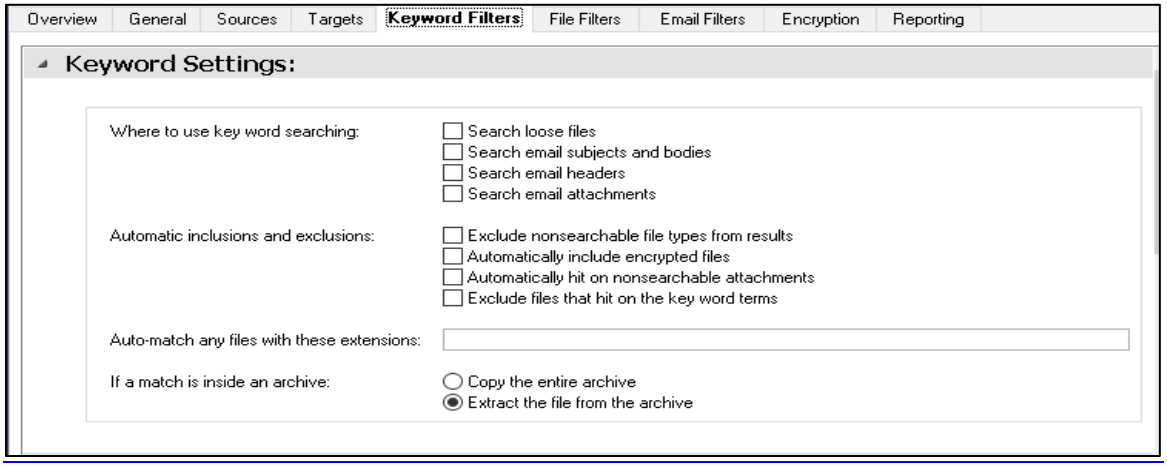
- **Do Nothing:**
  This ignores any files that already exist at the destination and does not include their counterparts in the source directory as responsive.
- **Overwrite Existing Files:**
  This option forces any files that already exist in the target folder to be overwritten.
- **Rename Files on Collision:**
  This option, when checked, will rename a file if a file by the same name already exists at the destination.

Keyword Filters Tab and File Filters Tab

KEYWORD FILTERS

[Harvester 5.1 Advanced Options - Part 3]

Keyword filtering is one of the most commonly used Harvester features. It is in this group of settings that users can perform targeted e-Discovery collections and filtering processes.



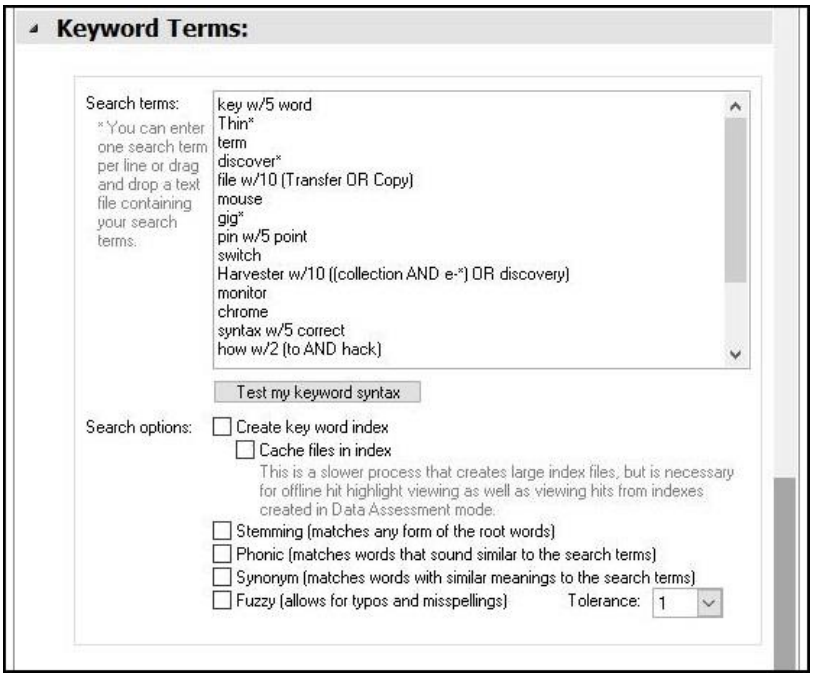
Harvester uses the superior search functionality provided by dtSearch. Many litigation support, computer forensics, and corporate IT professionals rely on dtSearch every day to rapidly and effectively search through large file collections.

- **Search loose files:**
  This option must be checked to enable key word filtering of what are commonly called *loose*, *native*, and *logical* files (i.e. Microsoft Word, Excel, PowerPoint, Acrobat PDF etc.).
- **Search email subjects and bodies:**
  This option must be checked to enable key word filtering of email subjects/bodies.
- **Search email headers:**
  This option must be checked to enable key word filtering of email headers.
- **Search email attachments:**
  This option must be checked to enable key word filtering of email attachments.
- **Exclude non-searchable file types from results:**
  When checked, this option excludes all file types that cannot be key word searched except those listed in the ***Exceptions*** box below. By checking the ***Exclude non-searchable file types*** box, you are instructing Harvester to exclude any files that are not considered keyword searchable based on their file type (executable, graphics, etc.).

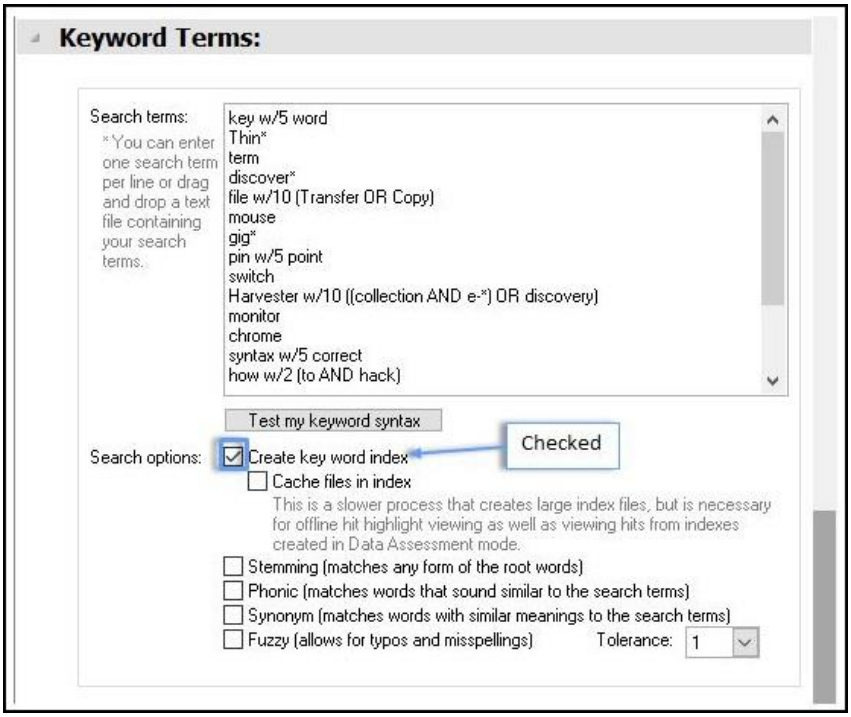
- **Automatically include encrypted files:**
  By checking the ***Automatically include encrypted files*** box, you are instructing Harvester to check to see whether any file of a type that can be encrypted (Office documents, PDF files, zip files, etc.) are, in fact encrypted before performing the keyword search and automatically issue a match for files that are encrypted (and also match all of the other non-keyword criteria).
- **Automatically hit on nonsearchable attachments:**
  By checking the ***Automatically hit on nonsearchable attachments*** box, you are instructing Harvester to consider any nonsearchable attachment (typically image files like jpeg or gif files) to be responsive and include the email in the results. This option is useful for collecting scanned documents of unknown format for later review. If you know the format, you can check the ***Exclude nonsearchable file types*** box and add the extension(s) that you wish to collect to the exceptions list below.
- **Archive Options (zip, rar, etc.):**
  If a key word hit appears in a file that is inside another (archive) file, you can either copy the entire archive file, or you can extract the file and create a folder structure on the target side named after the archive file that contained the hit as well as its internal folders.

INDEXING

Searching and copying data without indexing.
With neither Indexing option checked, Harvester will enumerate and copy all files with keyword hits but indexing and keyword hit highlighting will not be available. This is the fastest type of search.

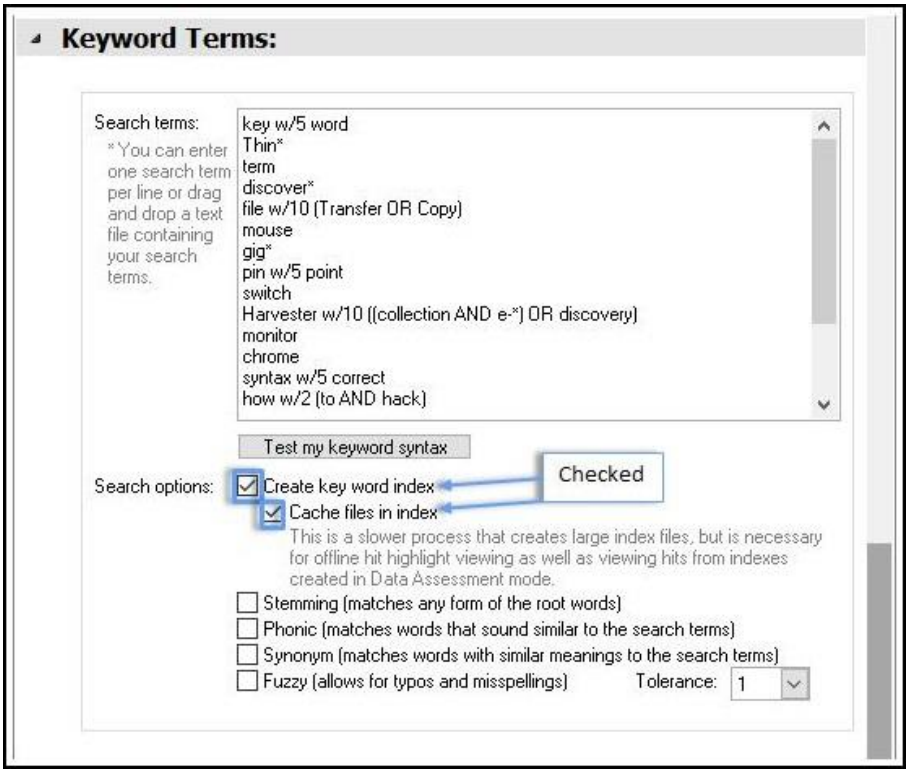


When *Create key word Index* option is selected on its own, Harvester will create the keyword index from the copied files at the end of the run in order to save time and to make the hit highlighting independent of the source.



When **Create Index** and **Data Assessment** options are selected Harvester will create a keyword index in the logs path during the enumeration phase. This allows you to generate keyword hit reports without actually copying any files.

Checking ***Create Cache Files in Index*** allows users to view offline document hit highlights.



NOTE: Creating full-text indexes (1&2) before the collection phase rather than from the collected information will take longer than when Harvester creates indexes from the collected information. The indexes will also be larger as a result of caching the file contents into the indexes.

After full-text indexes are created during Data Assessment Mode:

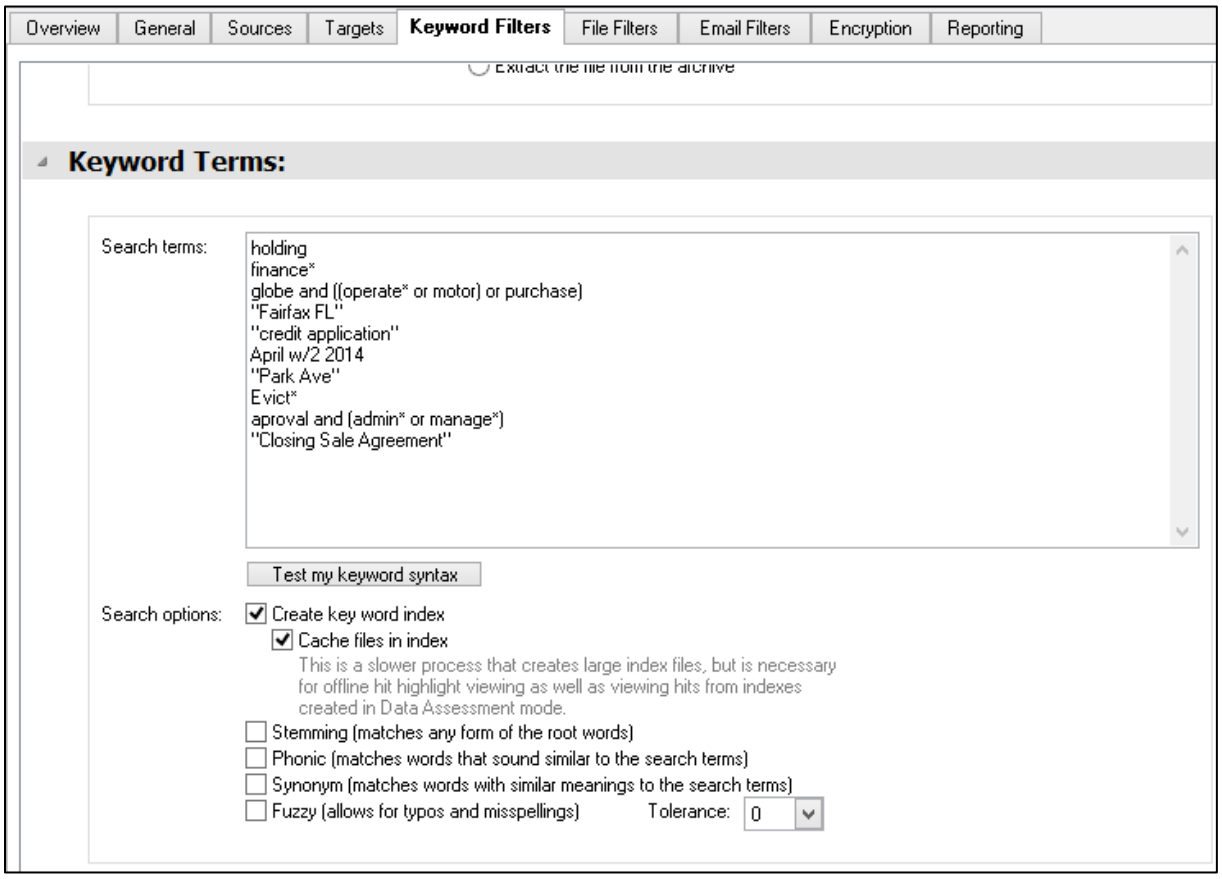
- Review keyword hit totals
- Create keyword hit reports
- View keyword hit highlighting for documents and email

For example, Harvester can be used to identify and index files from a remote system or server that may be offline or unavailable while using keyword hit highlighting. Caching file contents in the index enables users to view the information, without the copied data present.

NOTE: Indexing all data for a Custodian will reduce enumeration speeds, but can be beneficial during review.

SEARCH TERM KEY WORD SYNTAX

Users can enter search terms and phrases as shown using one term per line. Harvester will **OR** the terms, which flags items as a match if they are true for any of the individual conditions. Individual words, phrases and many other variations can be used as next outlined in the keyword syntax options.



<u>Document keyword search supports Boolean search requests:</u>

A Boolean search request consists of a group of words, phrases or macros linked by connectors such as *AND* and *OR* that indicate the relationship between them. Some examples include:

Search Request	**Meaning**
approval and management	both words must be present
approval or management	either word can be present
approval w/5 management	*Approval must occur within 5 words of management*
approval not w/12 management	Approval must occur, but not within 12 words of management
approval and not management	*Only approval must be present*
name contains smith	the field name must contain smith
approval w/5 xfirstword	*Approval must occur in the first five words*
approval w/5 xlastword	*Approval must occur in the last five words*

If you use more than one connector (and, or, contains, etc.), you should use parentheses to indicate precisely what you want to search for.

For example:
approval and ***management*** or ***withdrawn***
Could mean
(***approval*** *and* ***management***) or ***withdrawn***
Or
approval *and (****management*** *or* ***withdrawn****)*

For best results, always enclose expressions with connectors in parenthesis. An example is:
(Approval and Management) or (name contains Smith)

If you use more than one word as a search term, such as the name of a company or business (i.e. Apple Tree), you should use quotations to indicate precisely what you want to search for. For example, *Apple Tree* would become ***"Apple Tree"***.
NOTE: With the exception of special characters, punctuation is treated as a space.

Search terms may include the following special characters:

Character	Meaning
?	matches any character
=	matches any single digit
*****	matches any number of characters
%	fuzzy search
#	phonic search
~	stemming
&	synonym search
~~	numeric range

To enable fuzzy searching, phonic searching, synonym searching or stemming for all search terms, check their corresponding boxes.

Stemming: This option will find grammatical variations of the listed key words. A search for ***apply*** with this option checked would also find ***applies***, ***applying***, or ***application***.
NOTE: Checking this box will apply stemming to all terms in your list. If you need to apply stemming to only specific words in your list, add a tilde (**~**) after them in the key word list: ***apply~***

Phonic Search: This option will find words that sound like the key word terms you have listed. A phonic search for ***Smith*** would also return instances of ***Smythe***.
NOTE: Checking this box will apply phonic searching to all terms in your list. If you need to apply phonic searching only to specific words in your list, add a pound (**#**) character to them in the key word list: ***Smith#***

Synonym Search: This will search for word synonyms for any of your search terms using a comprehensive English language thesaurus or user-defined thesaurus terms. For instance, a synonym search for ***help*** would also return ***assist***.
NOTE: Checking this box will apply a synonym search to all terms in your list. If you need to apply synonym searching only to specific words in your list, add an ampersand (**&**) character after the word in the key word list: ***help&***

Fuzzy Searching: This option finds words even if they are misspelled. A search for *alphabet* with a fuzziness of 1 would also find ***alphaqet***. With a fuzziness of 3, the same search would find both ***alphaqet*** and ***alpkaqet*** . Fuzzy searching sifts through scanning and typographical errors. You can adjust the level of fuzziness from 1 to 10. (Usually values from 1 to 3 are best for moderate levels of error tolerance.)
NOTE: Checking this box will apply fuzzy searching to all terms in your key word list. If you need to apply fuzzy searching only to certain terms in your list, use the percent (**%**) sign within the word to indicate the first position where an error should be tolerated and repeat the sign for the number of errors that are tolerable from that point: ***a%lphabet*** would hit on ***alphaget*** and ***amphabet***. ***a%%%lphabe***t would hit on these as well as ***amphaket***.

Proximity Searches: Use the ***W/N*** connector in a search request to specify that one word or phrase must occur within N words of the other. For example, *approval w/5 management* would retrieve any document that contained approval within 5 words of management. The following are examples of search requests using W/N:

(approval or management) w/5 administrator
(approval w/5 administrator) w/10 management
(approval and administrator) w/10 management

Nested Searches
(this or that) w/10 (((work* and play*) or (sink w/2 hole)) or (quick w/1 sand))

This or **That** must be *within 10 words* of *both* **work** *and* **play**
or
This or **That** must be *within 10 words* of **sink** (which must be within 2 words of the word **hole**)
or
This or **That** must be *within 10 words* of **quick** (which must be within 1 word of the word **sand**)

Field Searches:

When indexing a database or other file containing fields, dtSearch saves the field information so you can perform searches limited to a particular field. For example, if you index an Access database with a *Name* field and a *Description* field, then you could search for “apple” in the Name field as below:

- **(Name contains apple)**

In addition to databases, dtSearch automatically recognizes metadata in supported file types. For a list of supported metadata formats, see "*What file formats does dtSearch support*" at http://support.dtsearch.com

Field searches can be combined using AND, OR, and NOT

- **(City contains (Portland or Seattle)) and (Address contains (Washington))**

The parenthesis is necessary to ensure that dtSearch interprets the search request correctly. More examples include:

- **(TO contains "example@email.com")**
- **(ATTACHMENTS contains "Test.xlsx")**

AUTOMATIC RECOGNITION SEARCHES:

Automatic recognition of dates:

Date recognition looks for anything that appears to be a date, using English-language months (including common abbreviations) and numerical formats. Examples of date formats that are recognized include:

- **January 15, 2006**
- **15 Jan 06**
- **2006/01/15**
- **1/15/06**
- **1-15-06**
- **The fifteenth of January, two thousand six**

To search for a date, put "date()" around the date expression or range. For example, to find any of the expressions above near the word "apple", search for:

- ***date(jan 15 2006) w/10 apple***

To search for a range of dates near the word "apple", search for:

- ***date(jan 10 2006 to jan 20 2006) w/10 apple***

Automatic recognition of email addresses:

Email address recognition looks for text that follows the syntax for a valid email address (example: *sales@pinpointlabs.com*). This makes it possible to search for a specific email address regardless of the alphabet settings for the "@" and "." characters, as well as any other punctuation that may be present in an email address.

To search for an email address, put "mail()" around the address. The “*” and “?” wildcard expressions are supported inside the () marks. Examples include:

- **mail(sales@pinpointlabs.com)**
- **mail(sa*@pinpointlabs.com)**

Automatic recognition of credit card numbers:

Credit card number recognition looks for any sequence of numbers that appears to satisfy the criteria for a valid credit card number issued by one of the major credit card issuers (MasterCard, Discover, Visa, etc.). Credit card numbers are recognized regardless of the pattern of spaces or punctuation embedded in the number. Examples include:

- ***1234-5678-1234-5678***
- ***1234567812345678***
- ***1234 5678 1234 5678***

Numerical tests used by the credit card issuers for card validity are used to exclude sequences of numbers that are not credit card numbers. However, these tests are not perfect and so the credit card number recognition feature may pick up some numbers that are not really credit card numbers.

To search for a credit card number, put "creditcard()" around the number.

- ***creditcard(1234*)***

Using keyword filters to extract from MBOX and DBX files:

Using dtSearch in Harvester, it is possible to extract EML files from MBOX and DBX mail stores. Extracting messages from MBOX and DBX files requires the use of keywords, as Harvester does not natively open and search these files. In the “Keyword Filters” tab, check “Search loose files” in the *Where to use key word searching* sections and select “Extract the file from the archive” in the *If a match is inside an archive:* section. To get messages matching certain criteria (date range, email addresses, etc.), use the automatic recognition searches listed above. To get all messages use the following term:

- ***a*, *e*, *i*, *o*, *u***

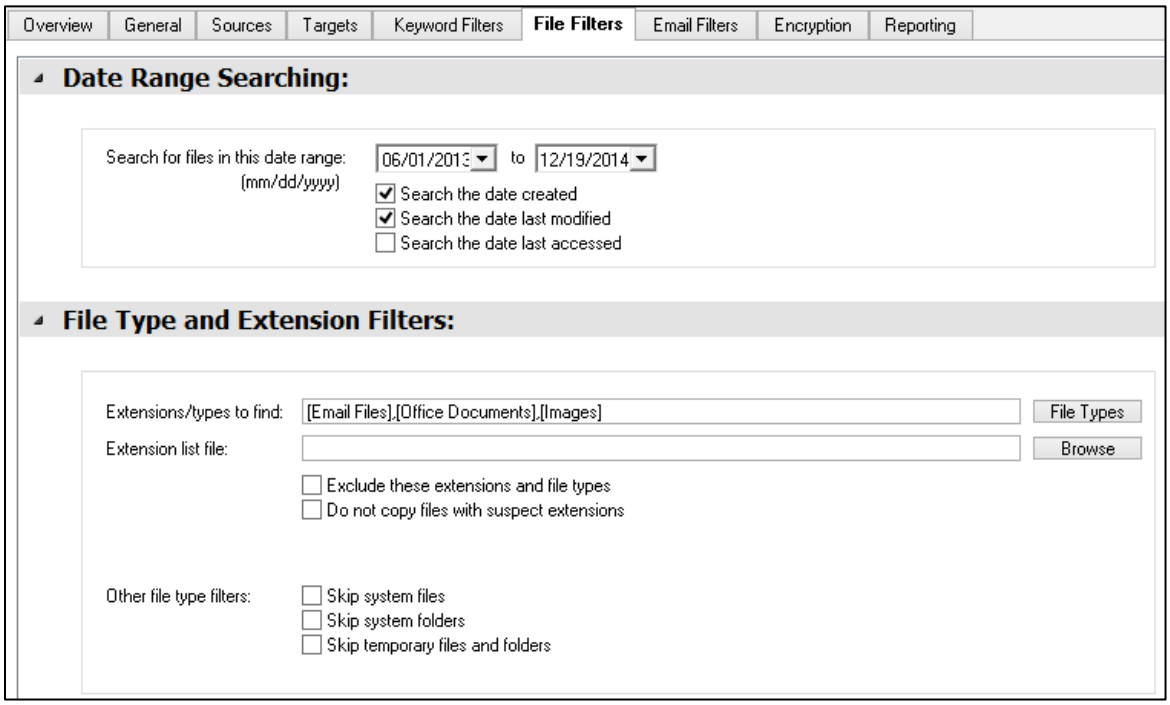
Test My Key Word Syntax:
Clicking on this button checks the syntax of the search terms entered. Errors will cause the search not to run. Warnings tell you that there is some ambiguity in the term and tell you how the search engine will assume you want the search run. If this matches your intentions, you can safely ignore the warning
NOTE: If there is a fatal error in the key word syntax, Harvester will prompt you with the

string(s) of syntax that are incorrect and warn you it will not be able to keyword search correctly with the error. For more information and how to fix the error, click the ***Test My Key Word Syntax*** button.

FILE FILTERS

Date Filter

Users can optionally add a date range filter for files. You can apply the date range to multiple time stamps by clicking the appropriate check boxes.



Created and modified times also apply to the archived files within a zip file if you have checked the ***Search zip files as directories*** box in the Sources tab.

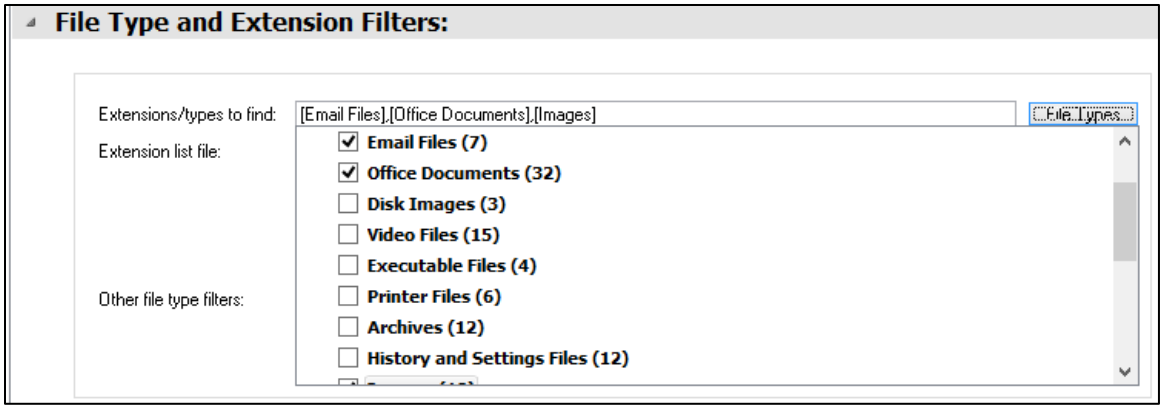
One or more boxes must be checked for the date range to apply.

- Creation Dates
- Last Modified Dates
- Last Accessed Dates

NOTE: This applies to loose files only – not to emails. Email date ranges may be set in the email option area.

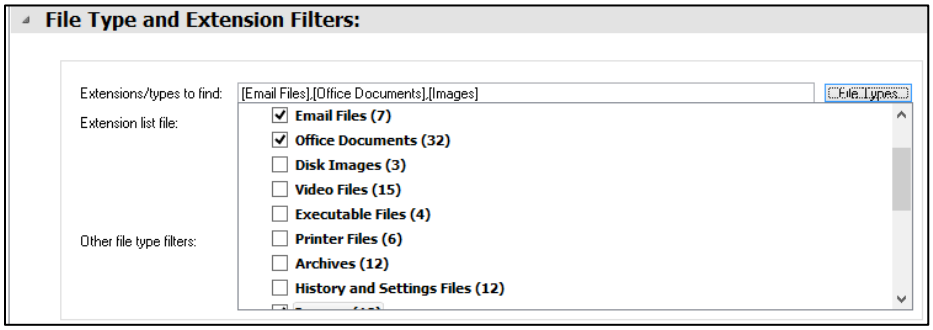
Extensions/Types to find

Harvester allows users to filter the data collected by file extension, file signatures, file type definitions and categories. Users can specify individual file extensions, file definitions (signatures) or categories, or create custom categories. Choose whether to include or exclude files with those file extensions in the dropdown box.



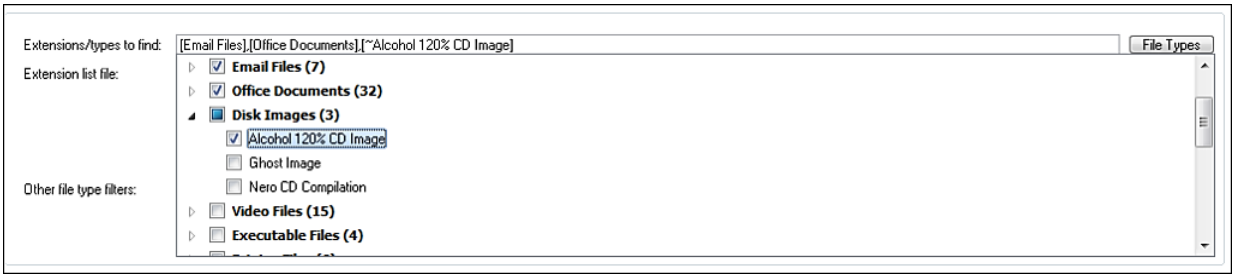
The following items can be specified:

- **File extensions – xls, xlsx, doc, docx, ppt, pptx, pdf, pst, ost, eml, msg** (specify with a comma delimiter). Users can also use file signatures (headers) instead of extensions by entering the name of the file definition in brackets with a tilde (~) character ([~MS Word] for example).
- **File Types – ([Office Documents],[Email Files],[Archives]).** If users would like to use file signatures (headers) or categories instead of extensions, they can click File Types and choose an entire category or click the dropdown icon for the individual file types.



NOTE: Files extensions without header signatures (ie: **.csv, .txt, .rtf**) must be added manually to be searched.

To select a category, click on the category and the ***Pick Category*** button.



This will automatically specify all files listed in the ***Individual File Types in Category*** column with a single [Category Name] reference. Choosing file types for header signature filtering will result in slower search speeds than file extensions alone because the file has to be opened during the search in order to read its header information.

Extension List File

Additionally, in the ***Extension List File*** section, you can use a text file containing multiple file extensions by browsing to the file using the button next to the field.

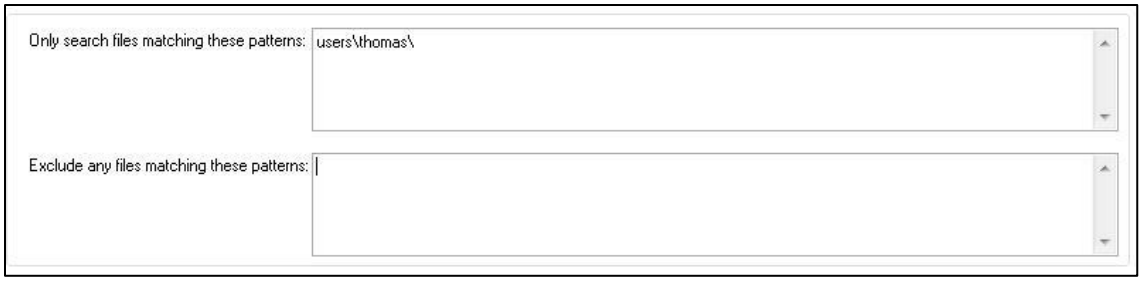
Exclude system files (with System attribute set): This option will filter out files which the file system (MFT/FAT) has flagged as system files. This is most commonly used in combination with the deNISTing option to further reduce the files collected.

Exclude system folders (System attributes set): This option will filter out folders (and included files) which the file system (MFT/FAT) has flagged as system folders. This is most commonly used in combination with the deNISTing option to further reduce the files collected.

Exclude temp files (with Temp attribute set): This option will filter out files which the file system (MFT/FAT) has flagged as temporary files. This is most commonly used in combination with the deNISTing option to further reduce the files collected.

ONLY SEARCH FILES MATCHING THESE PATTERNS

File Name inclusion options allow you to specify patterns that will be used to include only files or folders based on the names or patterns that you specify.

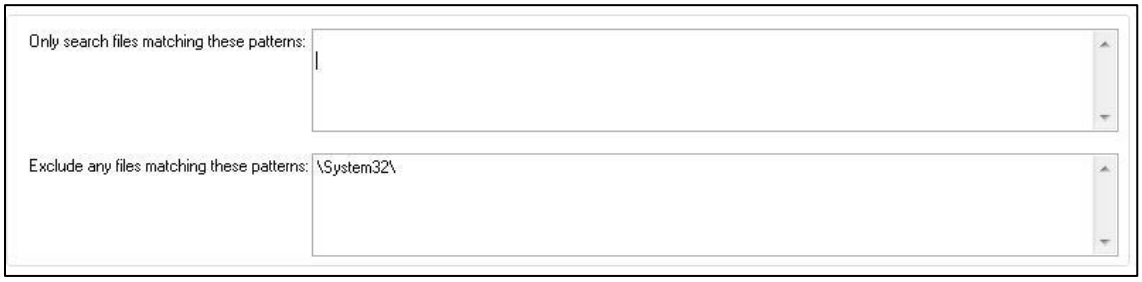


Multiple patterns can be added if needed, one per line. The syntax options are listed below.

Supported wildcard characters:
? – Any single character
***** - Zero or more characters
- Any single digit
[List of characters] – Any character in the list
[List of characters] – Any character not in the list
List syntaxes may contain either a simple list (**[1a7v]**) or a range indicator (**[0-9]** or **[a-f]**.
NOTE: These filters apply to whole paths. Comparisons are case insensitive. At time of comparison, all folders end with a **** character

EXCLUDE ANY FILES MATCHING THESE PATTERNS

Exclusion options allow you to specify patterns that will be used to exclude files or folders based on a mask.



Multiple patterns can be added if needed, one per line. The syntax options are listed below.

Supported wildcard characters:
? – Any single character
***** - Zero or more characters
- Any single digit
[List of characters] – Any character in the list
[List of characters] – Any character not in the list
List syntaxes may contain either a simple list (**[1a7v]**) or a range indicator (**[0-9]** or **[a-f]**.

NOTE: These filters apply to whole paths. Comparisons are case insensitive. At time of comparison, all folders end with a **** character

Deduping and Hash List Filtering

Exclude duplicates: This option filters out duplicate files within the current job. This process compares the MD5 hash value of each file and if a duplicate is identified, it will not be copied and an entry will be made in the exclusion log. It does not compare files within archives (i.e. ***Zip, RAR,TAR, Bzip, Gzip***) or mail stores. *An option to de-duplicate messages in Microsoft Outlook PST files is available under the Email options.*

Use Hash List Filter (deNIST): This option allows users to filter the source files against the NIST (National Institute of Standards and Technology) NSRL hash list and other included defined hash lists. The hash lists used for comparison are located in the \bin_hashlist directory. Any number of hash lists can be included. If a match is found in one of the hash lists, the file is logged along with the hash list that contained the matching hash.

Use Hash List Filter on emails: This option indicates that that the hash value that is listed in the NSRL or other defined hash lists will be used to filter emails in addition to loose documents.

Use Hash List Filtering on Email Attachments: This option indicates that that the hash value that is listed in the NSRL or other defined hash lists will be used to filter an email's attachments in addition to filtering loose documents.

Exclude Matching Hashes: This option indicates that files with a hash value that is listed in the NSRL or other defined hash list should NOT be copied.

Include Matching Hashes: This option indicates that files with a hash value that is listed in the NSRL or other defined hash lists are the ONLY files that will be copied.

HASH LISTS

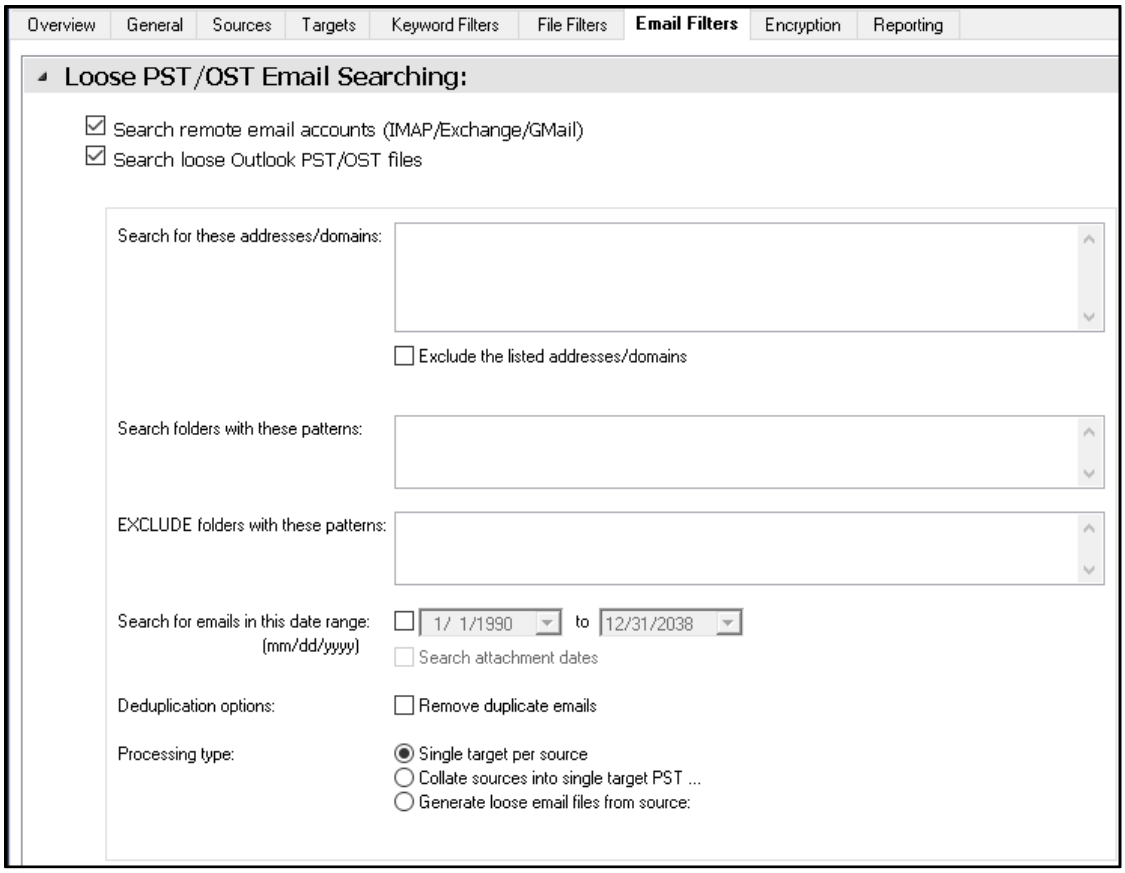
The ***Hash filtering and deNISTing*** options are useful in various applications.

- DeNIST using the MD5 NIST list to cull unnecessary data from your collection
- Add the hashlist from a previous job to the ***_hashlist*** folder and Exclude matching hashes for incremental backups
- Add the hashlist from a previous job to the ***_hashlist*** folder to copy the same data to another, or multiple computers using the ***Include matching hashes*** option
- Single or multiple hashlists can be used simultaneously in the ***_hashlist*** folder

Email Filters Tab

OUTLOOK EMAIL FILTERS

Email Filters allow users to search and copy messages from loose Outlook PST's and Exchange OST's without using Outlook or a MAPI connection. It also allows you to search email sources that are connected through Outlook via MAPI. When email sources are encountered during enumeration, Harvester automatically starts new threads to handle separate mail stores and reserves one thread to continue processing individual loose files. The number of threads is set to 'Auto' by default based on the system hardware; however, it can be customized by the user.



Selecting the corresponding box expands the options for each section.

PST EMAIL SEARCHING - SEARCH LOOSE OUTLOOK PST/OST FILES

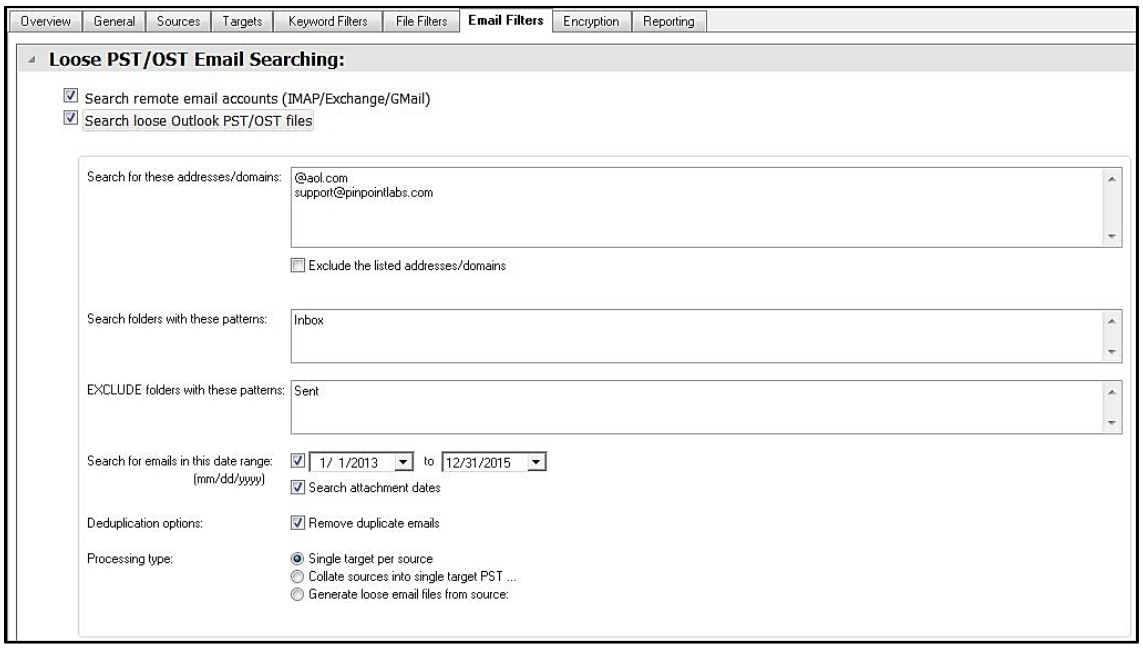
Harvester 5.1 Advanced Options - Part 3

Search for these addresses/domains (To/From/CC/BCC): You can enter or paste a list of items that are going to be used in the filter. There should be one entry per line as shown in the image. Names, domains or email addresses may be entered. When a domain only is entered, all emails from that domain will be selected.

Exclude emails with matching addresses: This option indicates whether emails that are found matching the ***Search for these addresses/domains*** (To/From/CC/BCC) filter should be included in the results or excluded from them.

Search folders with these patterns: This option allows you to enter the names of the folders in the PST to be searched. Use only one entry per line as shown in the image. Leaving this field blank will search all folders. This field supports the following wildcard characters:

- ***** matches any number of characters
- **?** matches any single character
- **#** matches any single digit



EXCLUDE folders with these patterns: This option allows you to enter the names of the folders in the PST that should NOT be searched.

This includes subfolders, so including ***SKIP_THIS_FOLDER*** in the exclusion patterns would skip any folder with ***SKIP_THIS_FOLDER*** (case insensitive) appearing in the path.

Both ***\SKIP_THIS_FOLDER*** and ***\Inbox\MyStuff\SKIP_THIS_FOLDER*** would be excluded. This field also supports the following wildcard characters:

- * matches any number of characters
- **?** matches any single character
- **#** matches any single digit

Start Date/Ending Date: These fields provide the option to narrow the emails extracted by the date range specified. This applies to emails only. The dates are entered in *MM-DD-YYYY* format.

Apply date range search to attachment file dates: Selecting this option applies the email date range filter to email attachments where applicable.

NOTE: Emails received via Exchange retain their original creation dates and modification dates, but attachments received via POP will have these dates set to the received time of the message.

Remove duplicate emails: When this option is checked, messages are compared across all PST files in the listed data sources. An MD5 hash value is calculated for each message and compared to all messages that have been processed in the current job. As duplicate messages are encountered they are flagged and written to the ***_duplicate_email.log***. The MD5 hash value is based on the following values: *Sender, Recipient, CC, BCC, Date, Subject, Email Body, Attachment Names, Attachment Sizes.*

Processing Type: This option determines the format for the target copies of the filtered messages.

- ***Create single target per source:*** This will create one target PST named the same as the original containing copies of the filtered messages. The new PST will reside in a path in the target according to the target path settings in **General Options**.
- ***Collate sources into a single target PST:*** This option will combine all source PSTs into the target specified in the **Process Target** path.
    - ***Process Target:*** Click the ***Browse*** button next to this field to specify the target PST. If no PST path is chosen, a PST file called ***collated.pst*** in the logs path will be used. This field supports the following variables:

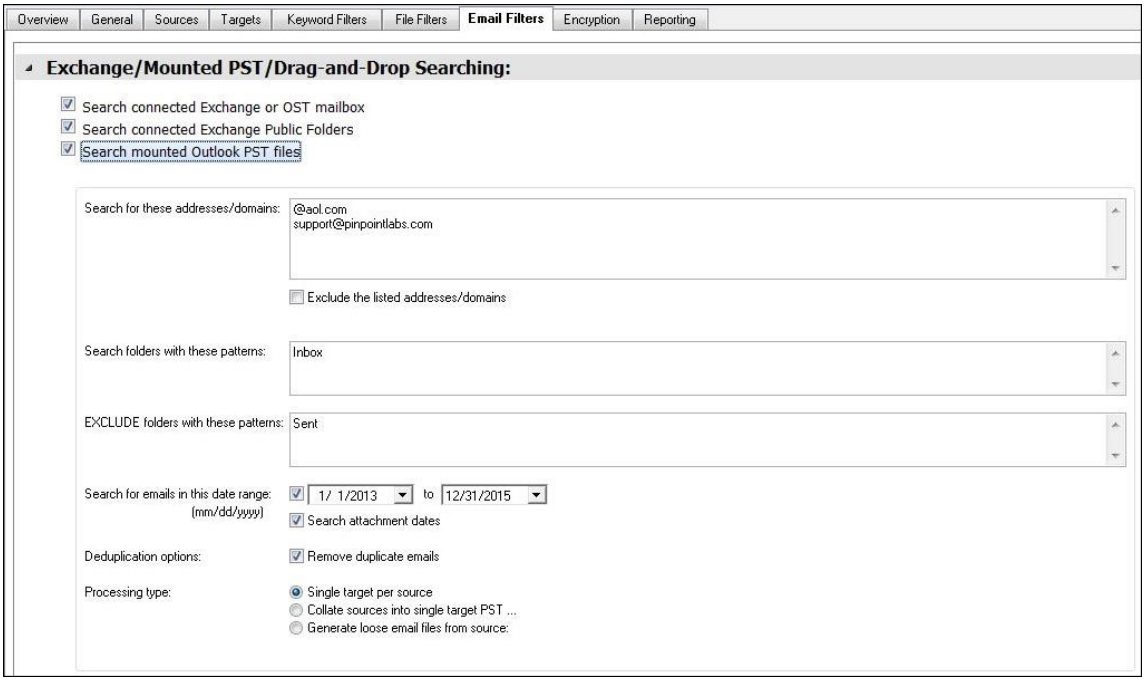
[SCDrive] – The drive letter that Harvester is running from.
[JobName] – The name of this job.
[Logs] – The path set up for logs.
[Target] – The path set up as the target for this job.
[CName] – The name of the computer running this job.
[UName] – The username of the logged-in user running this job.
[Date] – The date the job was run.
[DateTime] – The date and time (to the second) that the job was run.

- ***Generate loose email files from sources:*** This option allows you to export responsive emails to loose message files.

Export Type: This option allows you to specify the format for the extracted messages. A copy of each email matching the filtered criteria will be saved in the chosen format and the subject is used as the filename. The messages will be stored in the same folder structure from the PST and the parent level folder is named after the source PST. *Only **.msg** and **.eml** files will retain attachments.* The following loose message types are supported:

- **Unicode Outlook Message (msg) files**
- **Raw RFC822 (.eml) files**

EMAIL OPTIONS - MICROSOFT EXCHANGE/ACTIVE EMAIL/DRAG & DROP FILTERING



These options allow you to apply filtering to connected MS Exchange Mailboxes, PST files that are actively mounted in the user's Outlook, or connected Exchange Public Folders.

Search Connected Exchange Mailbox:
When checked, this searches and exports the resulting responsive messages from the default Exchange Mailbox connected to by the logged in user's Outlook.

Search Connected Exchange Public Folders:
When checked, this searches and exports the resulting responsive emails from Exchange Public Folders that are connected to the logged in user's Outlook.

Search Mounted MS Outlook PST Files:
When checked, this searches and exports the resulting responsive emails from mounted Outlook PST Files.

Search for these addresses/domains (To/From/CC/BCC):
You can enter or paste a list of items that are going to be used in the filter. There should be one entry per line as shown in the image. Names, domains or email addresses may be entered. When only a domain is entered, all emails from that domain will be selected.

Exclude the listed addresses/domains:
This option indicates that emails found matching the ***Search for these addresses/domains (To/From/CC/BCC)***filter should be excluded from the search results.

Folders to Search: This option allows you to enter the names of the folders in the PST to be searched. Use only one entry per line as shown in the image. Leaving this field blank will search all folders. This field supports the following wildcard characters:

- * matches any number of characters
- **?** matches any single character
- **#** matches any single digit

Folder Exclusion Patterns: This option allows you to enter the names of the folders in the PST that should NOT be searched. This includes subfolders, so including *SKIP_THIS_FOLDER* in the exclusion patterns would skip any folder with *SKIP_THIS_FOLDER* (case insensitive) appearing in the path. Both *SKIP_THIS_FOLDER* and *Inbox**MyStuff**SKIP_THIS_FOLDER* would be excluded. This field also supports the following wildcard characters:

- * matches any number of characters
- **?** matches any single character
- **#** matches any single digit

Start Date/Ending Date: These fields provide the option to narrow the emails extracted by the date range specified. This applies to emails only. The dates are entered in *MM-DD-YYYY* format.

Apply date range search to attachment file dates: Selecting this option applies the email date range filter to email attachments where applicable.

NOTE: Emails received via Exchange retain their original creation dates and modification dates, but attachments received via POP will have these dates set to the received time of the message.

Remove duplicate emails: When this option is checked, messages are compared across all PST files in the listed data sources. An MD5 hash value is calculated for each message and compared to all messages which have been processed in the current job. As duplicate messages are encountered they are flagged and written to the ***_duplicate_email.log***. The MD5 hash value

is based on the following values: *Sender, Recipient, CC, BCC, Date, Subject, Email Body, Attachment Names, and Attachment Sizes.*

Processing Type: This option determines the format for the target copies of the filtered messages.

Create single target per source: This will create one target PST named the same as the original containing copies of the filtered messages. The new PST will reside in a path in the target according to the target path settings in **General Options**.

Collate sources into a single target PST: This option will combine all source PSTs into the target specified in the ***Process Target*** path.

Process Target: Click the ***Browse*** button next to this field to specify the target PST. If no PST path is chosen, a PST file called ***collated.pst*** in the logs path will be used. This field supports the following variables:

[SCDrive] – The drive letter that Harvester is running from.
[JobName] – The name of this job.
[Logs] – The path set up for logs.
[Target] – The path set up as the target for this job.
[CName] – The name of the computer running this job.
[UName] – The username of the logged-in user running this job.
[Date] – The date the job was run.
[DateTime] – The date and time (to the second) that the job was run.

Generate loose email files from sources: This option allows you to export responsive emails to loose message files.

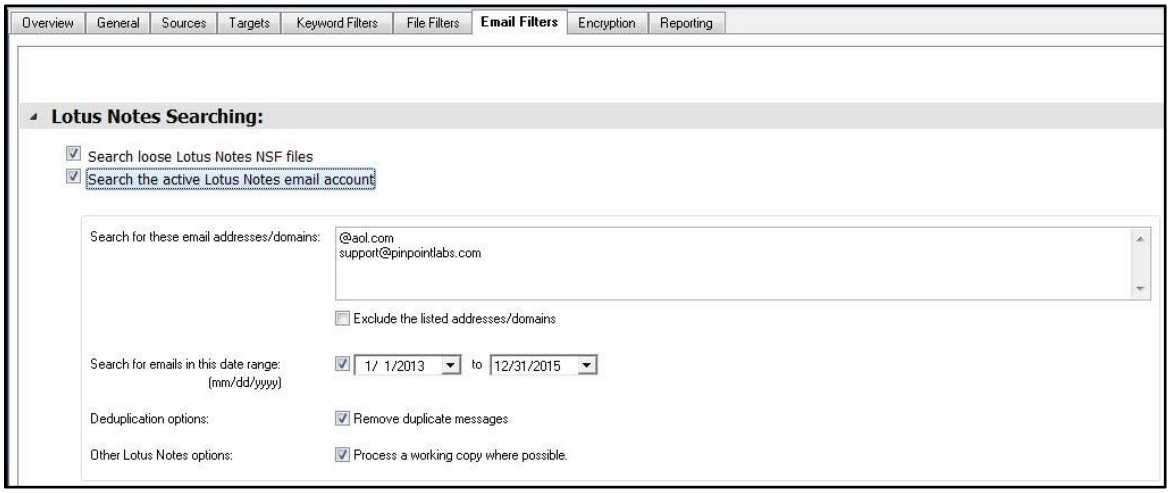
Export Type: This option allows you to specify the format for the extracted messages. A copy of each email matching the filtered criteria will be saved in the chosen format and the subject is used as the filename.

The messages will be stored in the same folder structure from the PST and the parent level folder is named after the source PST. Only ***.msg*** and ***.eml*** files will retain attachments. The following loose message types are supported:

- **Unicode Outlook Message (msg) files**
- **Raw RFC822 (.eml) files**

EMAIL OPTIONS - LOTUS NOTES

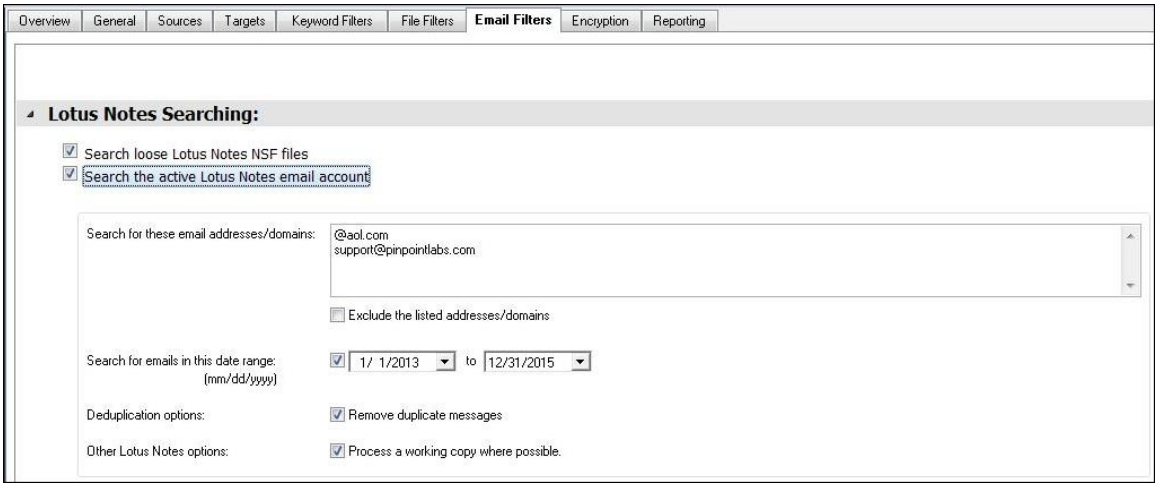
Email options allow users to filter Lotus Notes (NSF) files. Filtering criteria can be applied to the header (i.e. email addresses, domains and display name), subject, message body and attachments.



Search loose Lotus Notes NSF files:
This item must be checked in order to enable Lotus Notes NSF email processing.
NOTE: NSFs that are found with no messages matching the applied filters are written to the exclusion log if the **Exclusions Log** option in **Reporting** has been selected.

Search the Active Lotus Notes email Account:
This option instructs Harvester to connect to the default Lotus Notes mail store that is set up in the current user's profile and conduct the search on it. It can be used independently of the ***Search loose Lotus Notes*** NSF files option.

Search for these email addresses/domains (To/From/CC/BCC):
You can enter or paste a list of items that are going to be used in the filter. There should be one entry per line as shown in the image. Names, domains or email addresses may be entered. When a domain only is entered, all emails from that domain will be selected.



Exclude the listed addresses/domains:
This option indicates that emails found matching the ***Address/Domain to Search for (To/From/CC/BCC)*** filter should be excluded from the search results.

Start Date/Ending Date:
These fields provide the option to narrow the emails extracted by the date range specified. This applies to emails only. The dates are entered in MM-DD-YYYY format. The dates matched are the Send/Received times from the email header as well as the Lotus Notes document creation date.

Remove duplicate emails:
When this option is checked, messages are compared across all stored emails in the listed data sources. An MD5 hash value is calculated for each message and compared to all messages that have been processed in the current job. As duplicate messages are encountered, they are flagged and written to the ***_duplicate_email.log***. The MD5 hash value is based on the following values: Sender, Recipient, CC, BCC, Subject, Email Body, Attachment Names, and Attachment Sizes.

Search working copy of NSF where possible:
When this option is checked, Harvester will make a copy of the NSF file and search the copy. This is done because Lotus Notes is unable to open or search a read-only NSF file and as such will change the metadata on any NSF that it opens. Creating a working copy allows you to retain the original NSF's metadata and hash value and still conduct a Lotus Notes search. Unchecking this box will instruct Harvester to conduct the search on the NSF file in its original location.

***NOTE: If you are using email filtering in any of the above options as well as searching for keywords**, Harvester will treat the combination as an **AND** (any keywords will be searched for within responsive emails located using the email filters).*

ABOUT THE PINPOINT LABS MAIL PROCESSING ENGINE (PPLM)

The PPLM email processing engine was introduced in Harvester 5.0 and is used to process messages from all mail sources except Lotus Notes. PPLM is multi-threaded, and a sub folder will be created for each mail store that runs during a job.

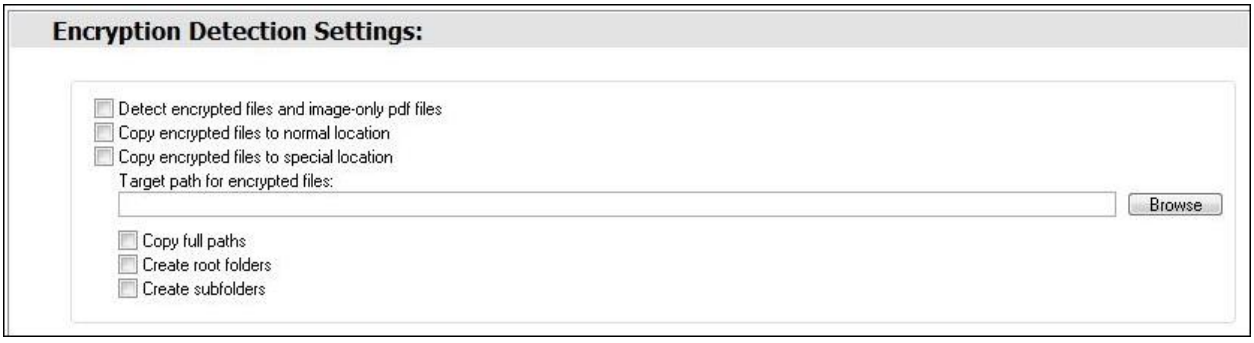
When troubleshooting, a Pinpoint Labs Support Engineer may ask for the logs in order to help diagnose a problem which can be found in the PPLM folder in your job log directory.

Encryption Tab and Reporting Tab

ENCRYPTED FILE DETECTION

[Harvester 5.1 Advanced Options - Part 3]

Harvester has the ability to identify several different types of encrypted files such as PST, PDF, Word, Excel, Access, and Zip files. The settings listed here let users determine whether to look for encrypted files and what to do with them if they are found.



Detect Encrypted and Image-Only Files: Checking this option will force all loose files and email attachments through the encryption detection routines. If these files are determined to be encrypted, they will be listed in the encrypted files log in the logs folder.

NOTE: In regards to PDF files, both the encryption status and whether a PDF file contains only images (*image-only*) are determined. *Image-only PDF files are considered encrypted because there is a high likelihood that they will need to be reviewed.*

Copy encrypted files normally: This option will copy encrypted files to their normal target locations. Unchecking this box instructs the program not to copy encrypted files to their normal target location.

Copy encrypted files to a special folder: This option allows the user to specify a target folder for any encrypted files. Users can either click the **Browse** button to select a folder or use the following variables to specify a target:

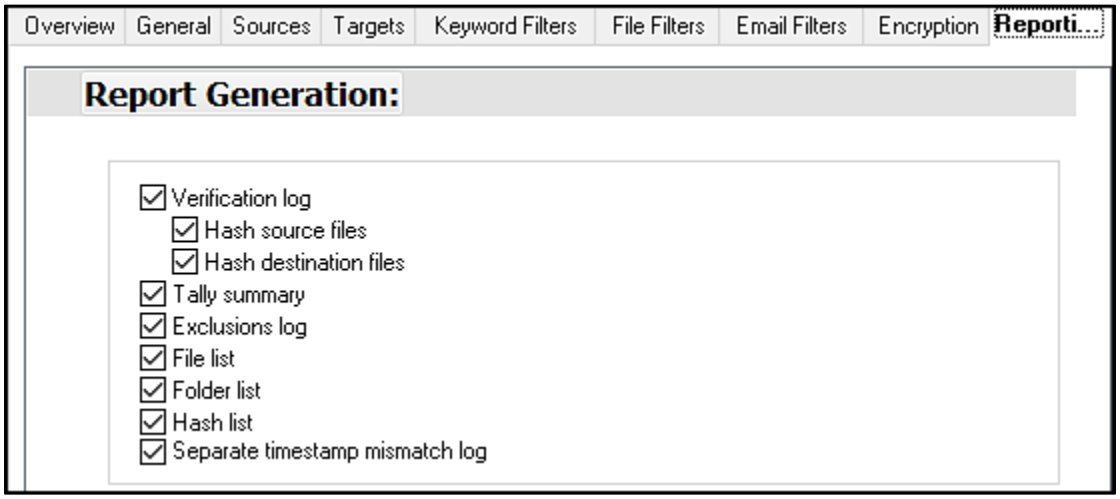
[SCDrive] – The drive letter that Harvester is running from.
[JobName] – The name of this job.
[Logs] – The path set up for logs.
[Target] – The path set up as the target for this job.
[CName] – The name of the computer running this job.
[UName] – The username of the logged-in user running this job.
[Date] – The date the job was run.
[DateTime] – The date and time (to the second) that the job was run

Copy Full Paths: This option will recreate the full path to the encrypted file on the target side.

Create Root Folders: This option will also create a folder at the base of the target path named after the drive letter or UNC server on which the file was found. For example, an encrypted file found at *C:\demo\test3\Crypto.doc* may be copied to *J:\Collected Files\Encrypted\C\demo\test3\Crypto.doc.*

Create subfolders: This option will create subfolders beneath the encrypted file target. Not checking this box or the *Create Full Paths* box will force all of the encrypted files into the root of the path specified above.

REPORTING



Verification Log: When selected, this option will create the Chain of Custody log file (_verification_log.csv). This report, a comma separated values file (.csv), lists fields pertinent to the copies made and the statistics of each file.

Using the ***Create Verification Log***, ***Hash the Source File*** and ***Hash the Destination File*** options will result in a detailed Chain of Custody log file saved in the directory chosen in the *Log File Path*.

These fields include:

- **Date/Time Copied**
- **Hashes Match**
- **TS Exact Match**
- **Source Path**
- **Source Created Date**
- **Source Modified Date**
- **Source Access Date**
- **Source Size (in bytes)**
- **Source MD5 (calculated MD5 hash value)**
- **Destination Path**
- **Destination Created Date**
- **Destination Modified Date**
- **Destination Access Date**
- **Destination Size (in bytes)**
- **Destination MD5 (calculated MD5 hash value)**
- **Error Messages**

Hash the Source File:
This option calculates the MD5 hash value of each file copied before the copy is made. The values are reported in the Chain of Custody log file.

Hash the Destination File:
This option calculates the MD5 hash value of each file once copied to the destination. The values are reported in the Chain of Custody log file.

File List:
This option stores a file containing the path and file name of each responsive file encountered for this job. It will also create individual extended file, extended email, extended archived files and extended email attachments lists. These lists can be created without copying the files.

Folder List:
This option stores a file in the specified log path that contains the top-level folders specified as sources.

Hash List:
This option writes the MD5 hash value for all responsive files to a hash list file located in the current

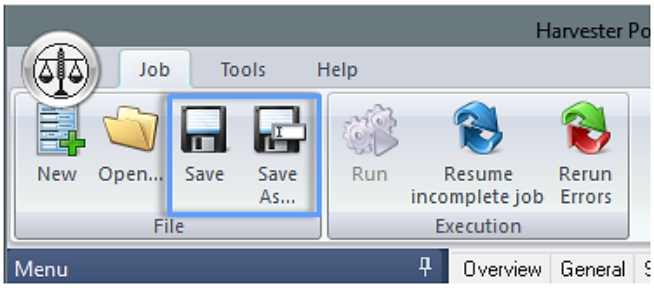
job log directory. This list can then be used as a filter (de-dupe) using the ***Use Hash List Filter (deNIST)*** option by placing this file in the _hashlist folder.

Exclusion Log:
This option creates a log of any files that were excluded due to the various exclusion filters or due to the *Hash List Filter* or ***Exclude Duplicates*** options. The log also contains an explanation for the exclusion.

Log Time Stamp Changes Separately:
When selected, the program will not count time stamp discrepancies due to mismatched file systems as errors in the error log, but will create a separate log to note these discrepancies.

SAVING HARVESTER JOBS

Once a Harvester job is created, users can save it to use with automated collections or reuse when needed. Harvester files are saved with the ***.occ*** extension.



Selecting ***Save As*** will open the ***_occ*** directory in the Harvester directory. Job files stored in the _occ directory will be displayed in the job list automatically when Harvester is launched.

Harvester job files can be quickly created, saved and emailed to clients for self-collection and stored wherever a user prefers. The Harvester job filename will default to the ***Job Name*** entered under details; however, users can choose an alternative.

OPENING HARVESTER JOBS

Harvester job files stored in the _occ directory will be displayed in the Job Profile tab when the application is launched. However, both job files and runtime .scj files can be loaded from other locations by clicking ***Open*** on the ***Job*** ribbon bar and browsing to the file location. Users can edit job files and update the file by clicking ***Save***.

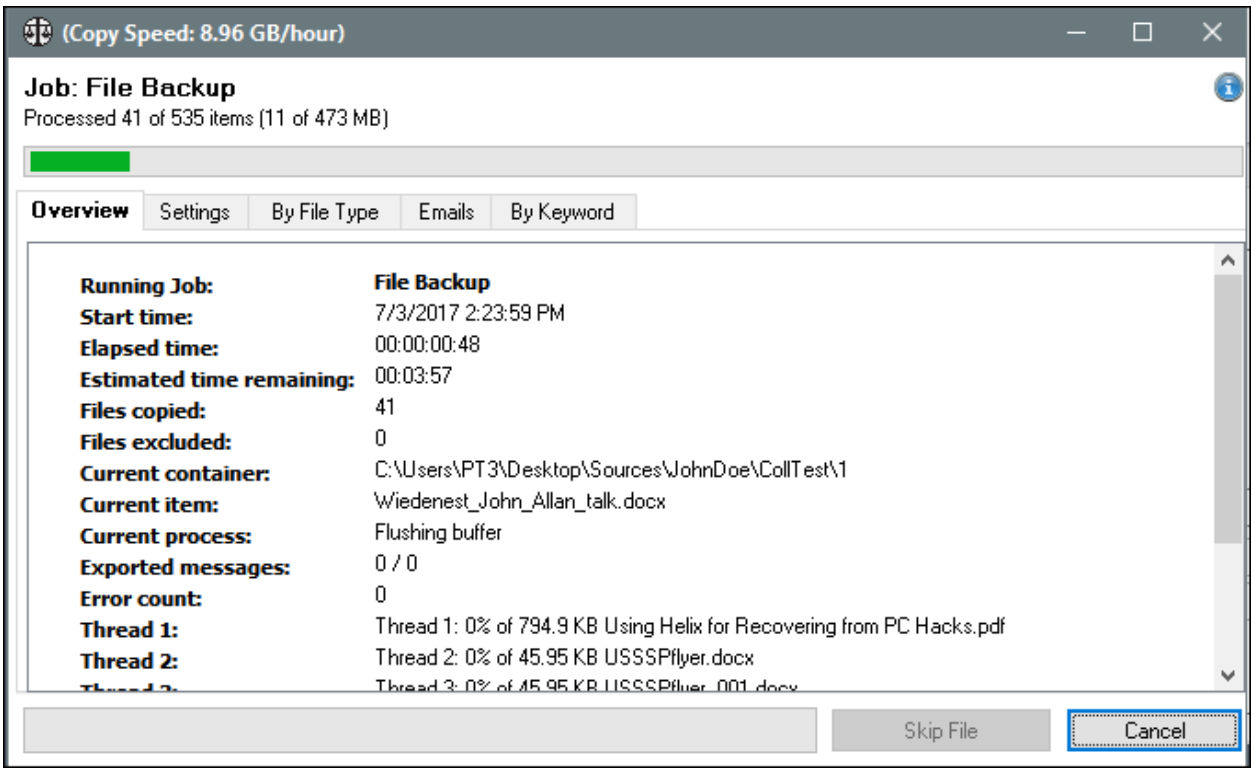
Users can also drag and drop an occ file from Windows Explorer to the job profile tree to display and edit settings. Clicking ***Run*** on the settings form will execute the current specifications in the job manager form.

Job Console

JOB CONSOLE OVERVIEW

Reference Video: [Running Harvester Server Locally]

The Harvester progress console will provide important feedback as well as valuable real-time statistics



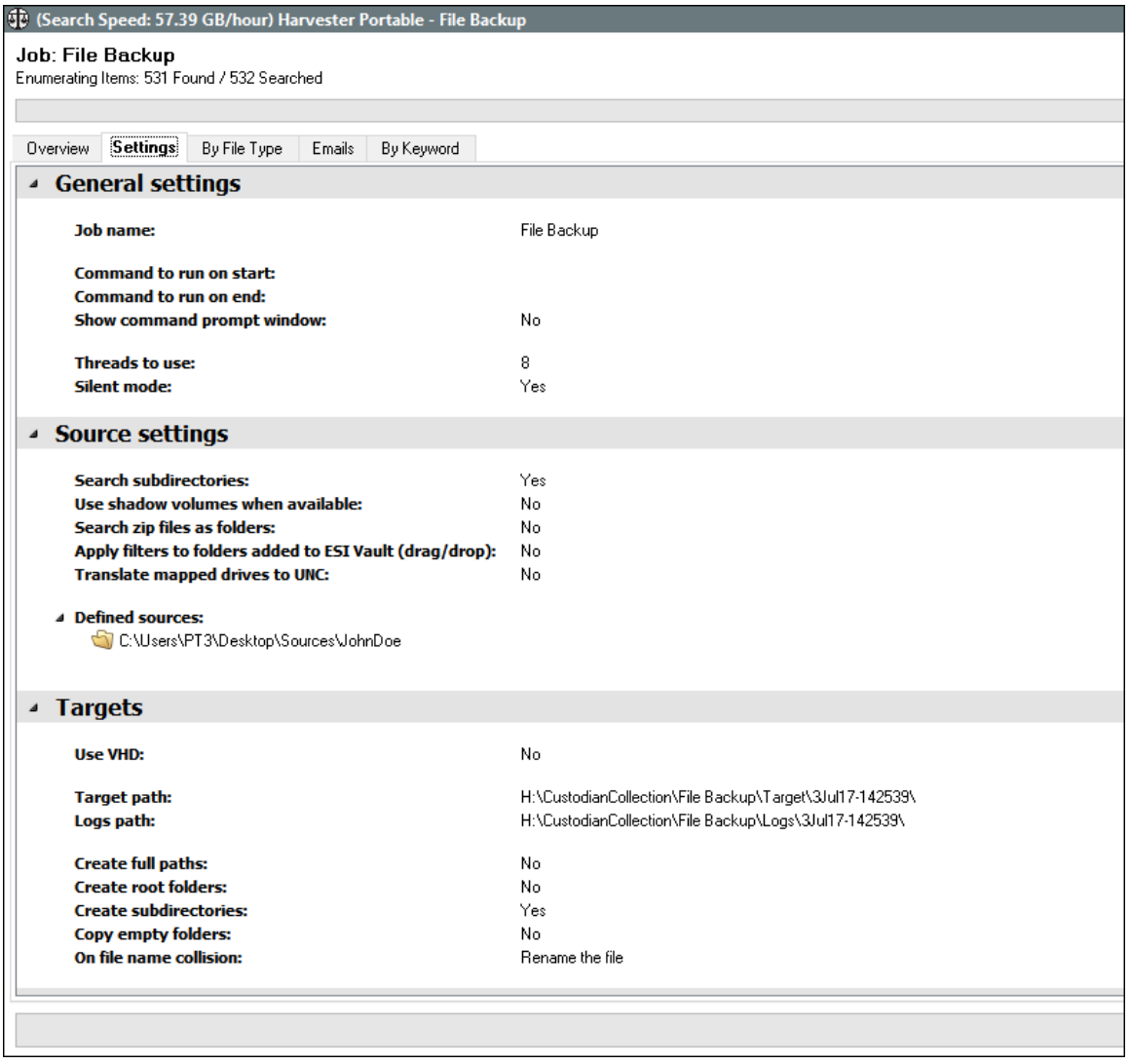
- Job Name
- Start Time
- Elapsed Time
- Estimated time remaining (during the processing phase)
- File included/excluded
- Current container
- Multi-threaded object identification

Users frequently monitor this tab for overall job processing information.

JOB CONSOLE SETTINGS

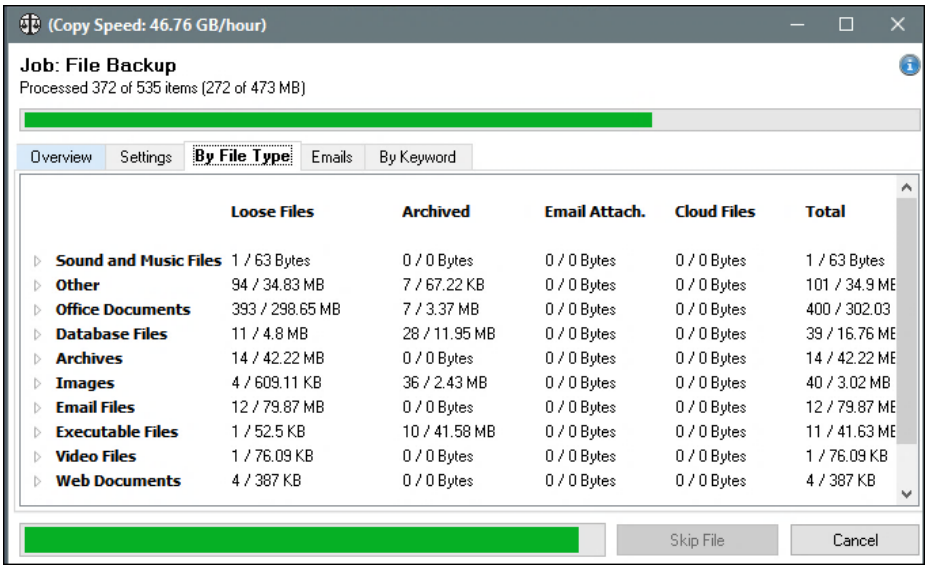
When running jobs directly from a profile on the local system (versus remote launching jobs on target computers) the Harvester job progress console will appear. Users can view their job profile settings while it's running by clicking on the **Settings** tab. Double clicking on the ***Target path*** or ***Logs path*** will open the corresponding locations.

Users can view their job profile settings while it's running by clicking on the ***Settings*** tab. Double clicking on the ***Target path*** or ***Logs path*** will open the corresponding locations.



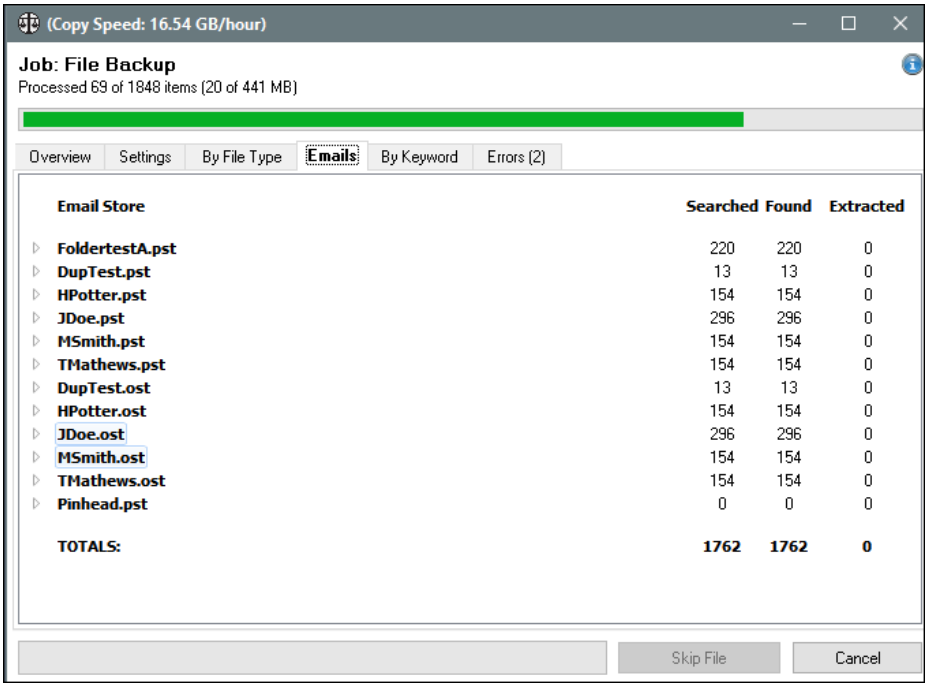
JOB CONSOLE BY FILE TYPE

While processing a job the ***By File Type*** tab will provide real-time statistics for file types and categories including total count and size. The matching files are also organized by ***Loose Files***, ***Archived***, and ***Email Attachments***.



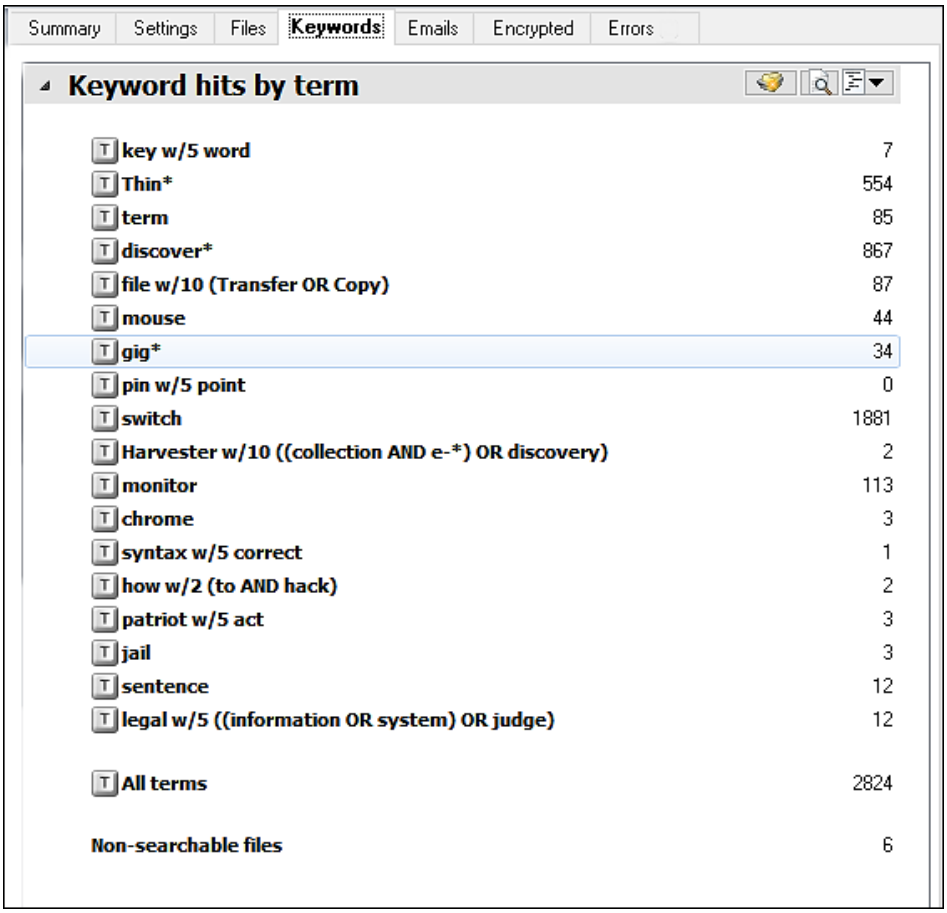
JOB CONSOLE EMAILS

While processing a job, the **Emails** tab provides real-time statistics for matching messages. The path to the store and total count are also included.



JOB CONSOLE BY KEYWORD

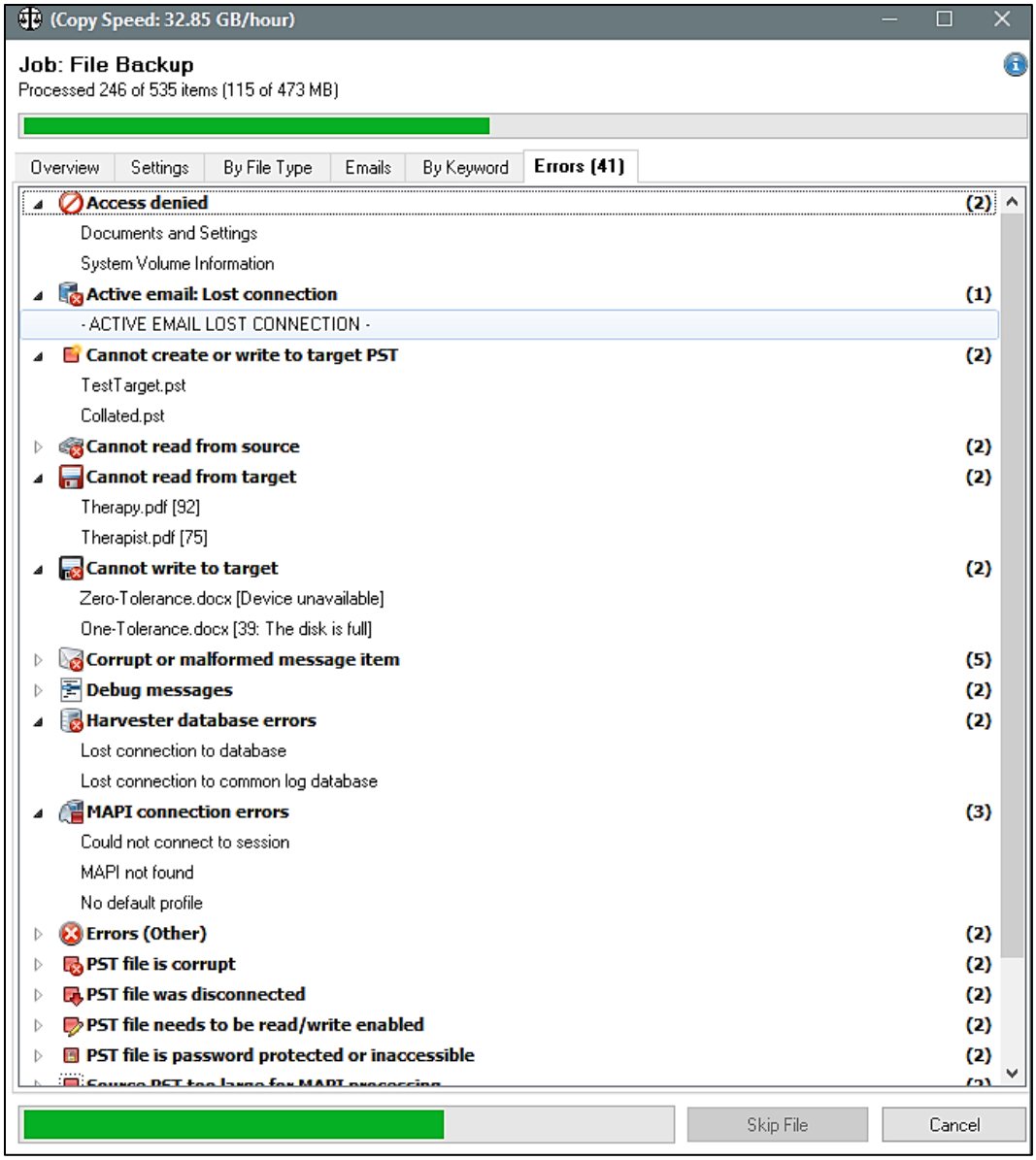
While a job is running, a list of the keywords in the job profile will be displayed in the **By Keyword** tab. Once the indexing process is completed, the counts will be updated. The **Create Index option must be selected to get keyword result counts.**



JOB CONSOLE ERRORS

While a job is running, users can see real-time error reporting on the Errors tab. Messages will be organized into common categories and by expanding a selection, users can see details related to each item.

- File-based errors allow you to double-click on the error to open an explorer window to the specified file.
- PST-based errors allow you to double-click on the error and run the source PST through ScanPST.



When running jobs directly from a profile on the local system (versus remote launching jobs on target computers) the Harvester job progress console will appear. Users can view their job profile settings while it's running by clicking on the **Settings** tab. Double clicking on the ***Target path*** or ***Logs path*** will open the corresponding locations.

RESUMING A JOB

A number of conditions may cause jobs to be cancelled or fail to complete. Incomplete jobs can be easily resumed by selecting the **Resume an Incomplete Job** option in the **Job Profile** ribbon. To resume the job, browse and select the _jobfile.scj, which will be located in the ***Logs Path*** that was specified for the project.

Alternatively, users can resume a job listed in the **Previous Jobs** list by clicking on the job and choosing **Resume this job**. This option can save time as users do not need to confirm the log folder location and manually browse to locate the .scj file.

RERUNNING ERRORS FOR A JOB

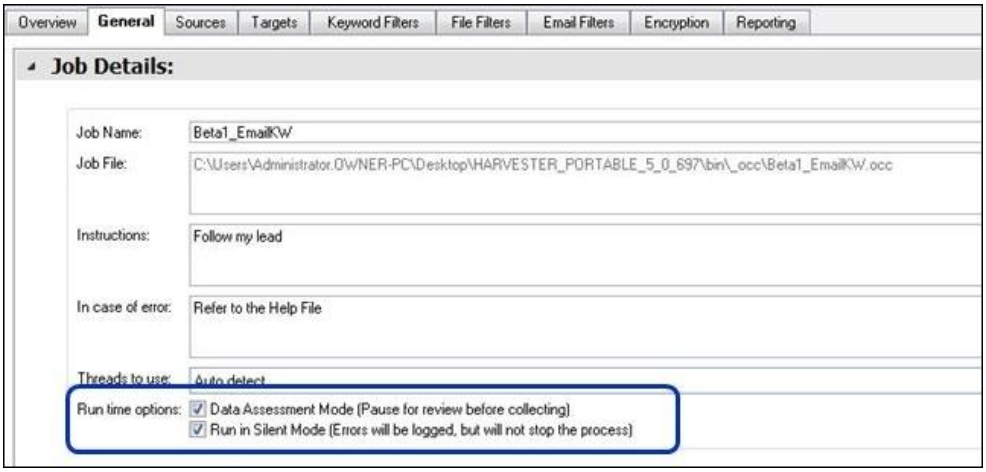
Jobs can also error out on certain files while being processed. Files with errors can be rerun simply by selecting the **Resume an Incomplete Job** option in the **Job Profile** ribbon. To rerun errors, you need to also browse and find the _jobfile.scj as you would if you were trying to resume an incomplete job.

Alternatively, users can rerun errors for jobs listed in the **Previous Jobs** list by clicking on the job and choosing **Rerun errors**. This option can save time as users do not need to confirm the log folder location

Data Assessment Mode

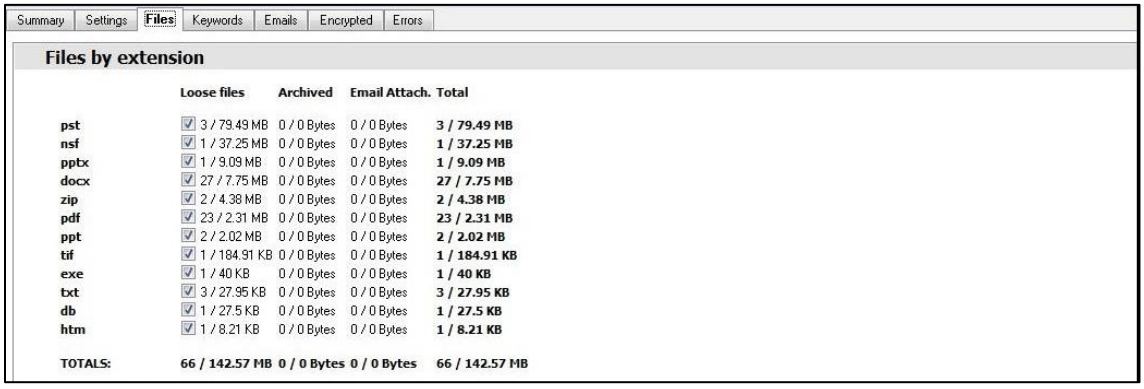
Harvester runs in two separate stages:
Enumeration (inventory stage): This is the first phase where it goes through the specified sources and records which files meet filter criteria.
Processing (copy stage): Harvester goes through the list it made during enumeration, and copies the items and hash verifies the copies.

Running Harvester in ***Data Assessment Mode*** will stop the job after enumeration, to generate inventory reports without copying the data.



NOTE: Clicking ***Resume this job*** at any point will finish enumeration and proceed to copy phase.

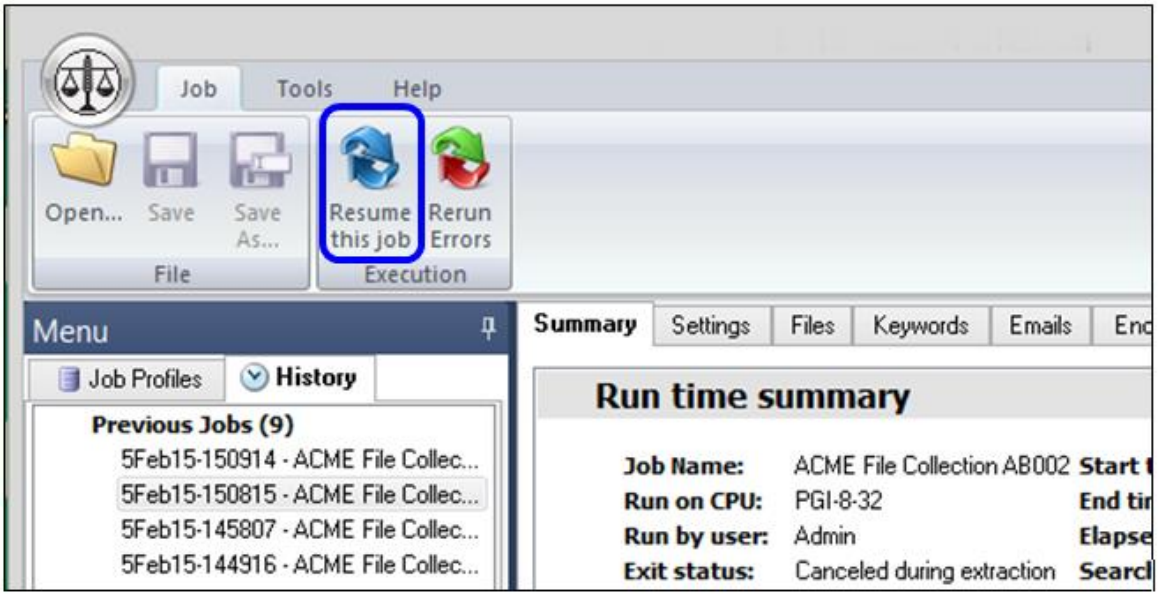
Post-Data Assessment is viewed in the History window under the **Files tab**. The user has the opportunity to uncheck unnecessary extensions before resuming the copy phase. This can cull the data further, potentially reducing copy time and per-gigabyte processing fees.



	Loose files	Archived	Email Attach.	Total
pst	3 / 79.49 MB	0 / 0 Bytes	0 / 0 Bytes	**3 / 79.49 MB**
nsf	1 / 37.25 MB	0 / 0 Bytes	0 / 0 Bytes	**1 / 37.25 MB**
pptx	1 / 9.09 MB	0 / 0 Bytes	0 / 0 Bytes	**1 / 9.09 MB**
docx	27 / 7.75 MB	0 / 0 Bytes	0 / 0 Bytes	**27 / 7.75 MB**
zip	2 / 4.38 MB	0 / 0 Bytes	0 / 0 Bytes	**2 / 4.38 MB**
pdf	23 / 2.31 MB	0 / 0 Bytes	0 / 0 Bytes	**23 / 2.31 MB**
ppt	2 / 2.02 MB	0 / 0 Bytes	0 / 0 Bytes	**2 / 2.02 MB**
tif	1 / 184.91 KB	0 / 0 Bytes	0 / 0 Bytes	**1 / 184.91 KB**
exe	1 / 40 KB	0 / 0 Bytes	0 / 0 Bytes	**1 / 40 KB**
txt	3 / 27.95 KB	0 / 0 Bytes	0 / 0 Bytes	**3 / 27.95 KB**
db	1 / 27.5 KB	0 / 0 Bytes	0 / 0 Bytes	**1 / 27.5 KB**
htm	1 / 8.21 KB	0 / 0 Bytes	0 / 0 Bytes	**1 / 8.21 KB**
TOTALS:	**66 / 142.57 MB**	**0 / 0 Bytes**	**0 / 0 Bytes**	**66 / 142.57 MB**

Under the Files tab, check for unnecessary file extensions. Choose and uncheck unnecessary extensions before resuming into the copy phase.

In the History Tab, highlight the job and click on ***Resume this job*** after review to proceed to copy phase.



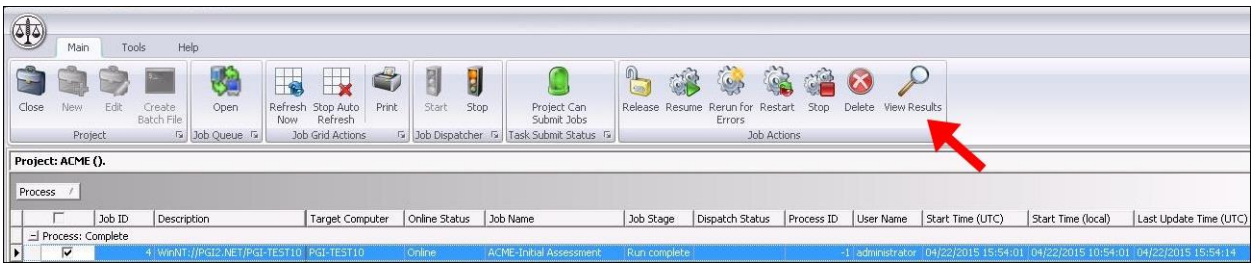
To Resume a job at a later time (Collect the identified data).

If the logs are available but the job is not listed under the History tab, open the Harvester, Click on Tools > Resume an Incomplete Job. Browse to the ***_jobfile.scj*** for the job you wish to resume and click ***Open***. The *_jobfile.scj* will be located in the logs folder.

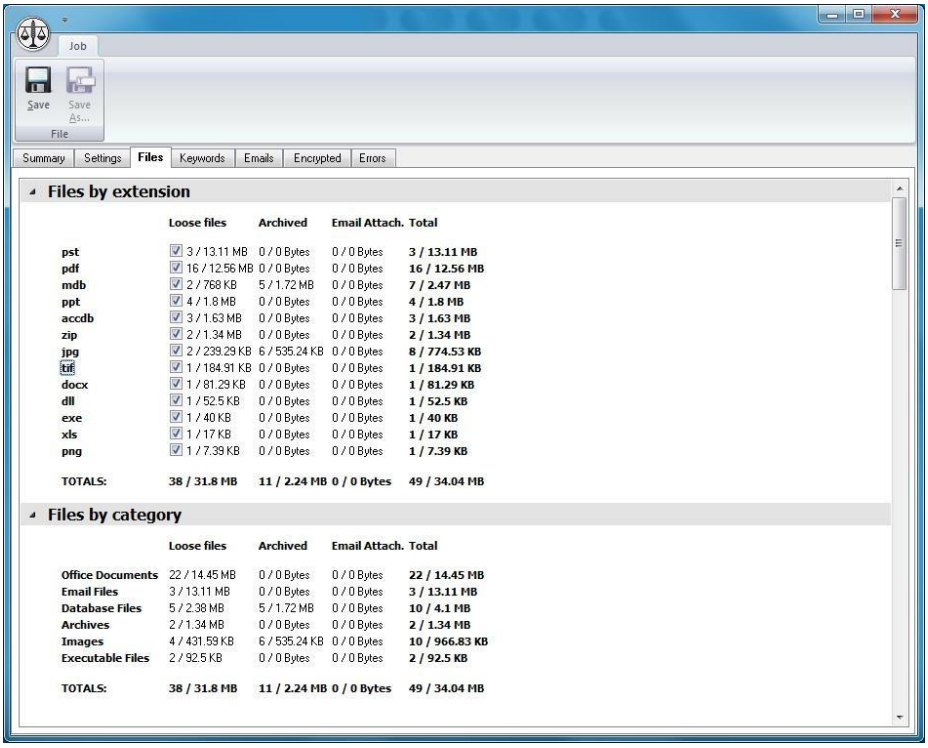
Post-Data Assessment is viewed in the History window under the **Files tab**. The user has the opportunity to uncheck unnecessary extensions before resuming the copy phase. This can cull the data further, potentially reducing copy time and per-gigabyte processing fees.

To Resume a Job via Harvester Server Console Interface (Collect the identified data)

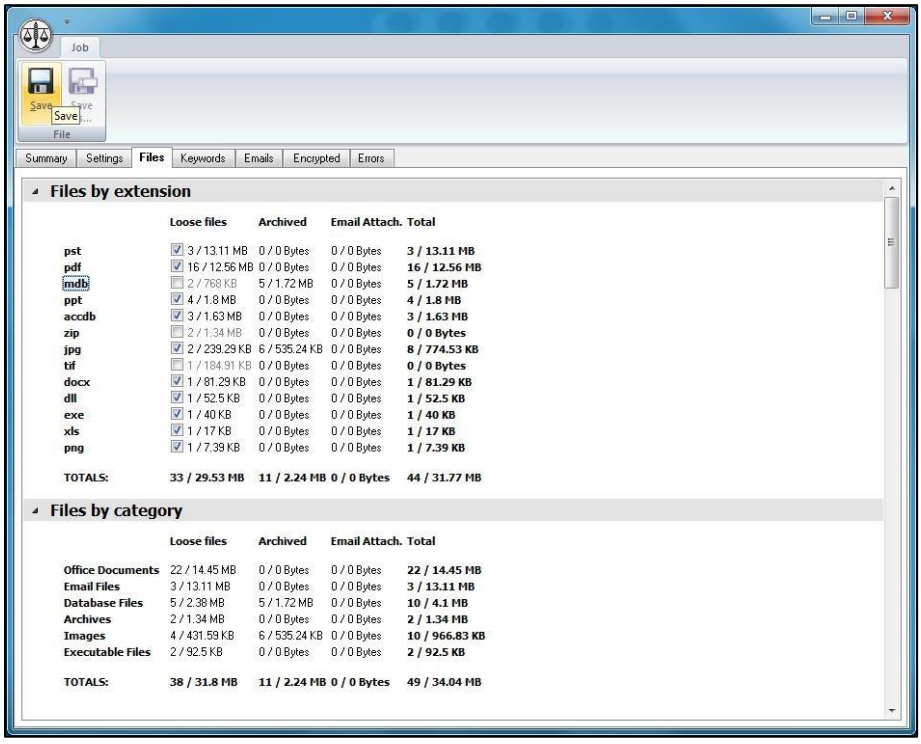
Highlight a job by checking its box and a Click the View Results button.



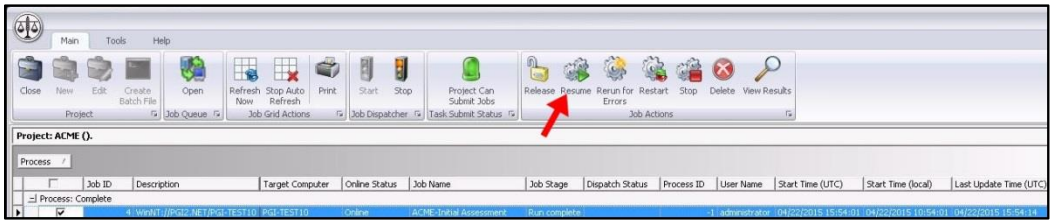
Under the Files tab, check for unnecessary file extensions.



Choose and uncheck unnecessary extensions before resuming into the copy phase.



Click Save, close the window and then click Resume.



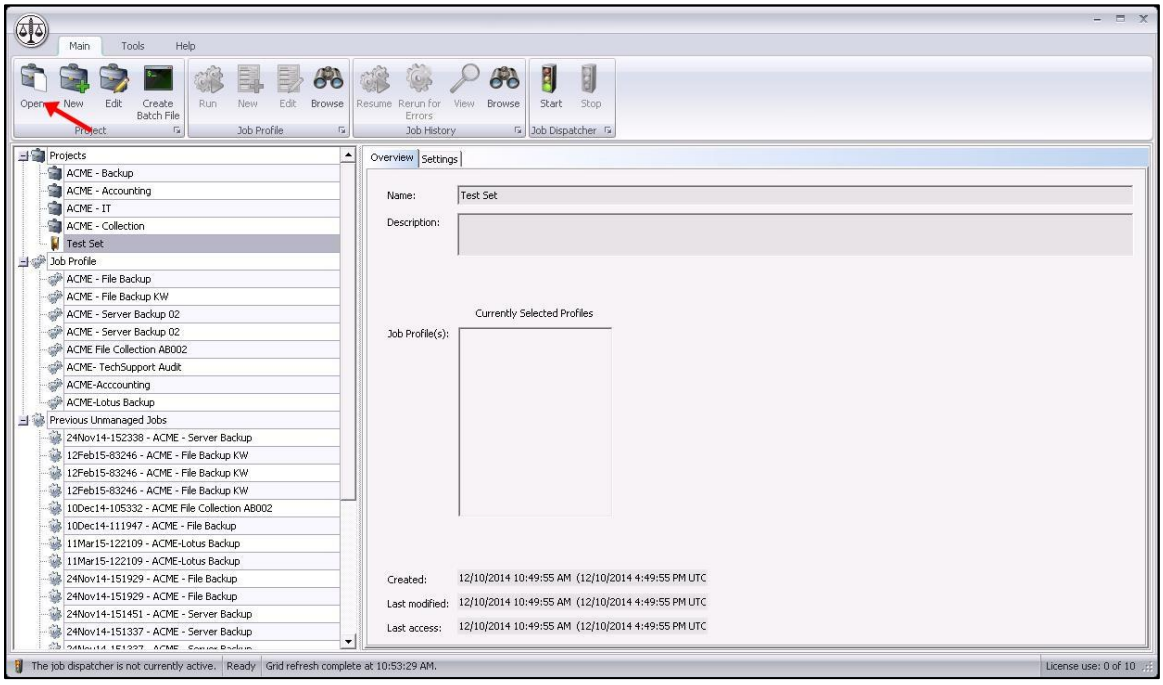
Job History

Harvester Server provides quick access to previously run job statistics and settings. A history database (History.db) is located in the Harvester Server application folder and will store the location and overall statistics of each completed job that is run from that Harvester executable.

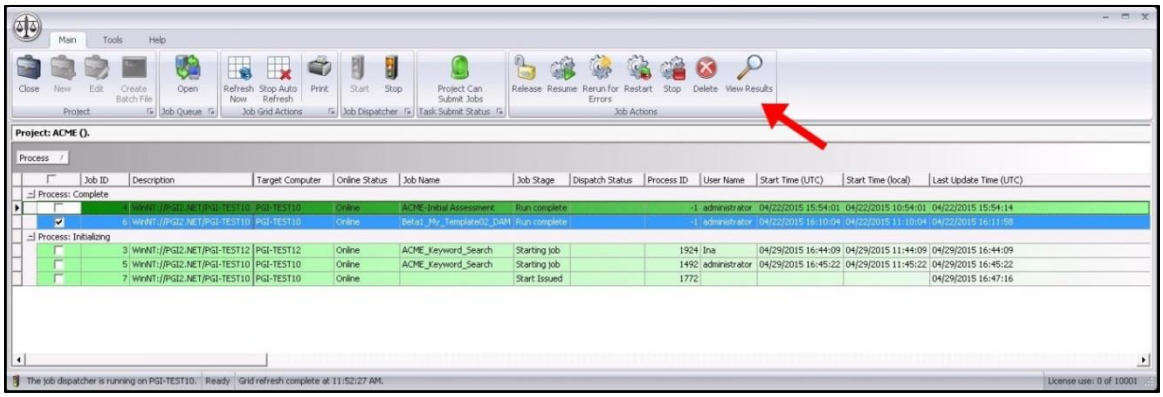
After a Harvester job completes, the progress bar will disappear and highlight the job history file that displays the ending job statistics and other useful details. If users create a new install from the Harvester archive file, then a new history database will be created.

TO VIEW PREVIOUS *PROJECT MANAGED* JOBS:

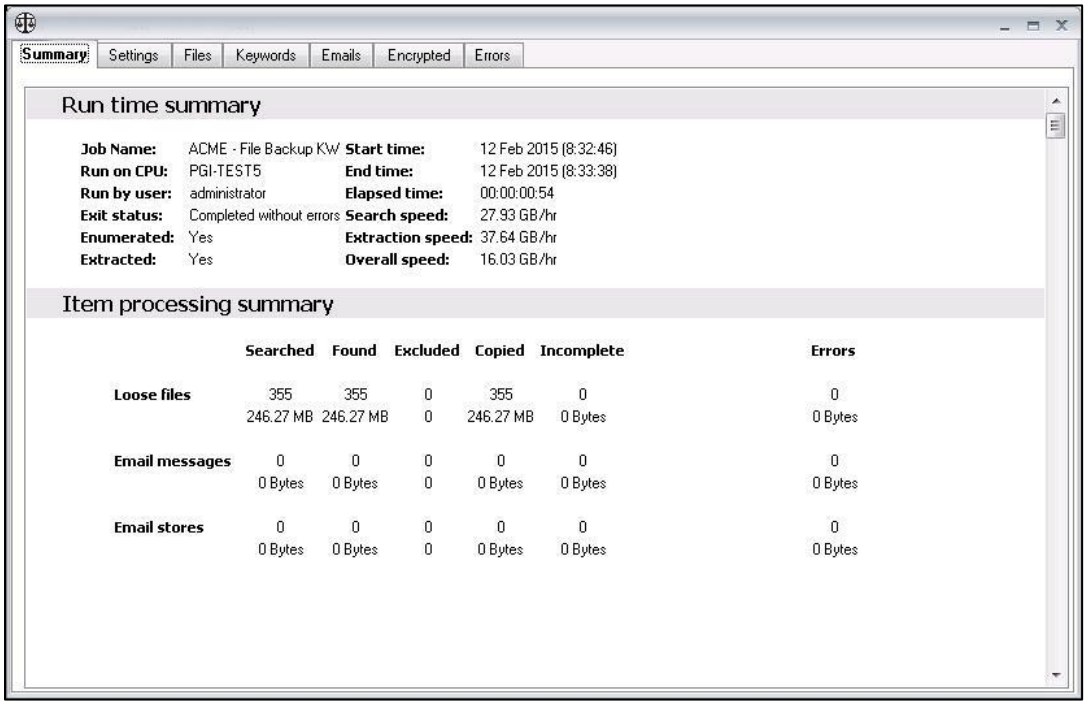
Highlight the project with the Profile you would like to view, and click ***Open***.



Check (highlight) a job and click ***View Results*** 

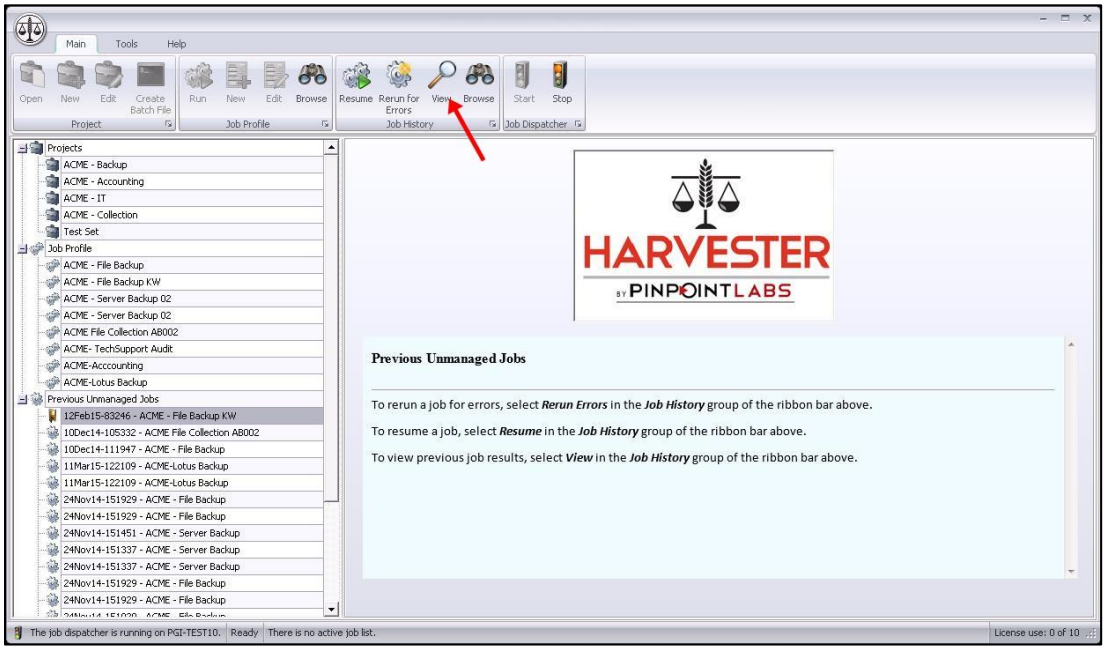


This will open the profile's history display interface (below):

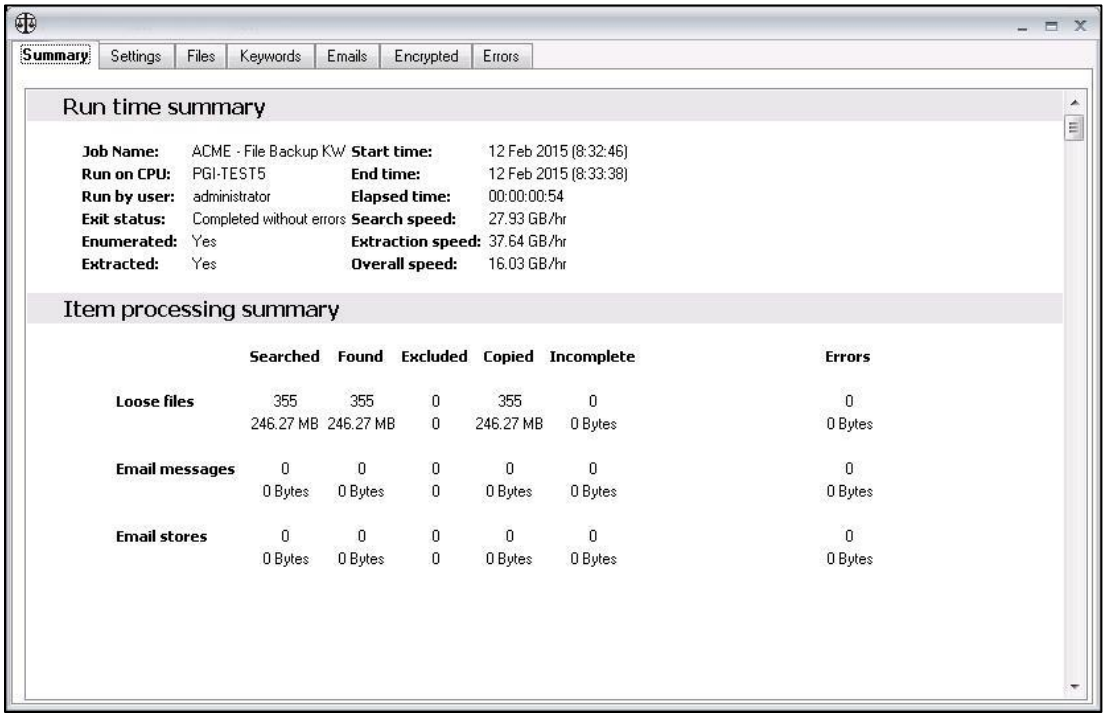


TO VIEW PREVIOUS UNMANAGED JOBS

Highlight the unmanaged job you would like to view and click the ***View*** button.

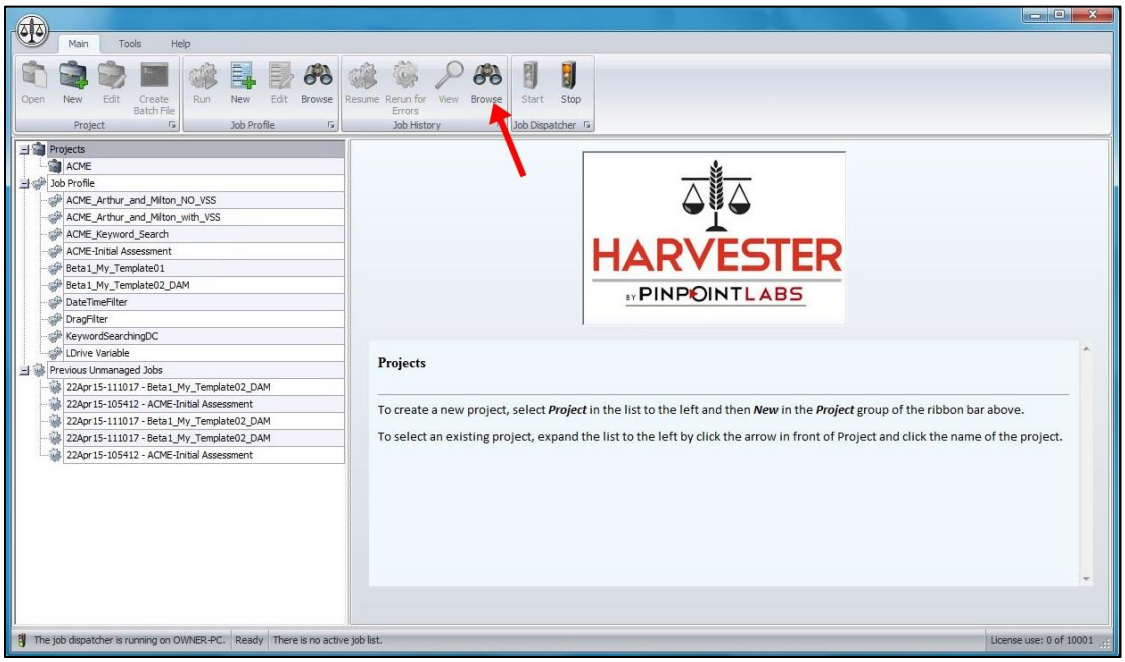


This will open the Job's History display interface (below):

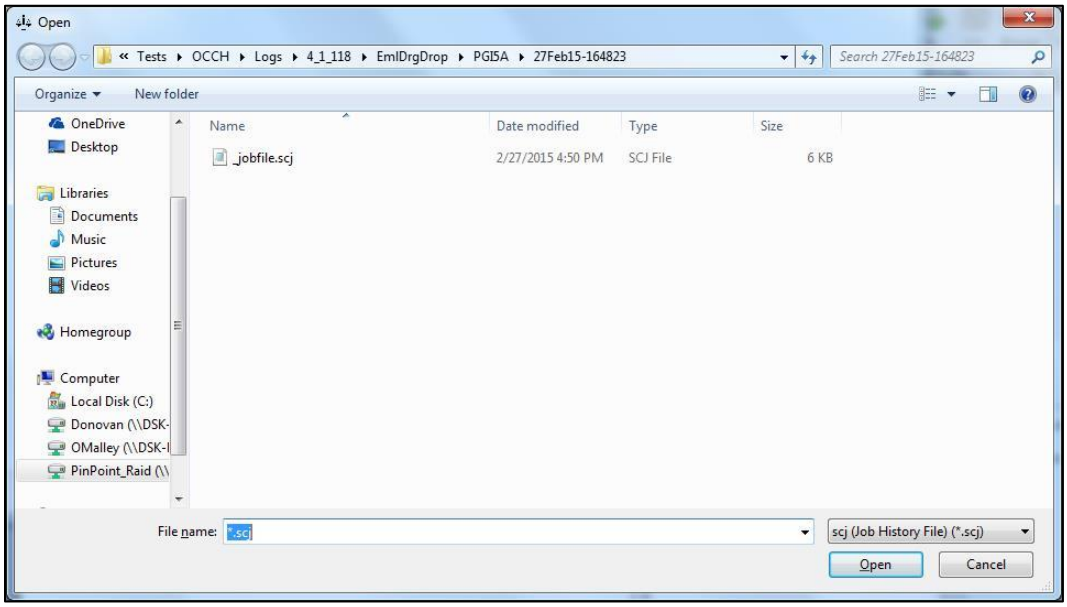


Harvester Server job files can also be opened with other versions or installs of Harvester.

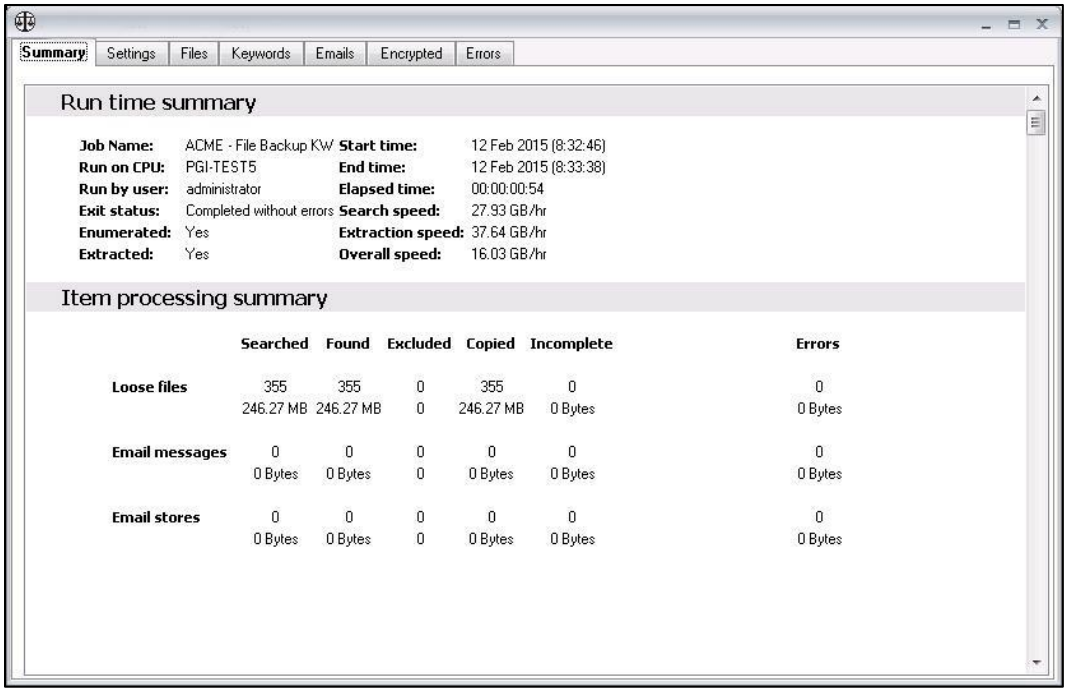
Click the Browse button in Job History



Browse to the specific logs folder and choose the ***_jobfile.scj***



This will open the Job's History display interface (below):



Harvester will display the job view and load the job details and results that include:

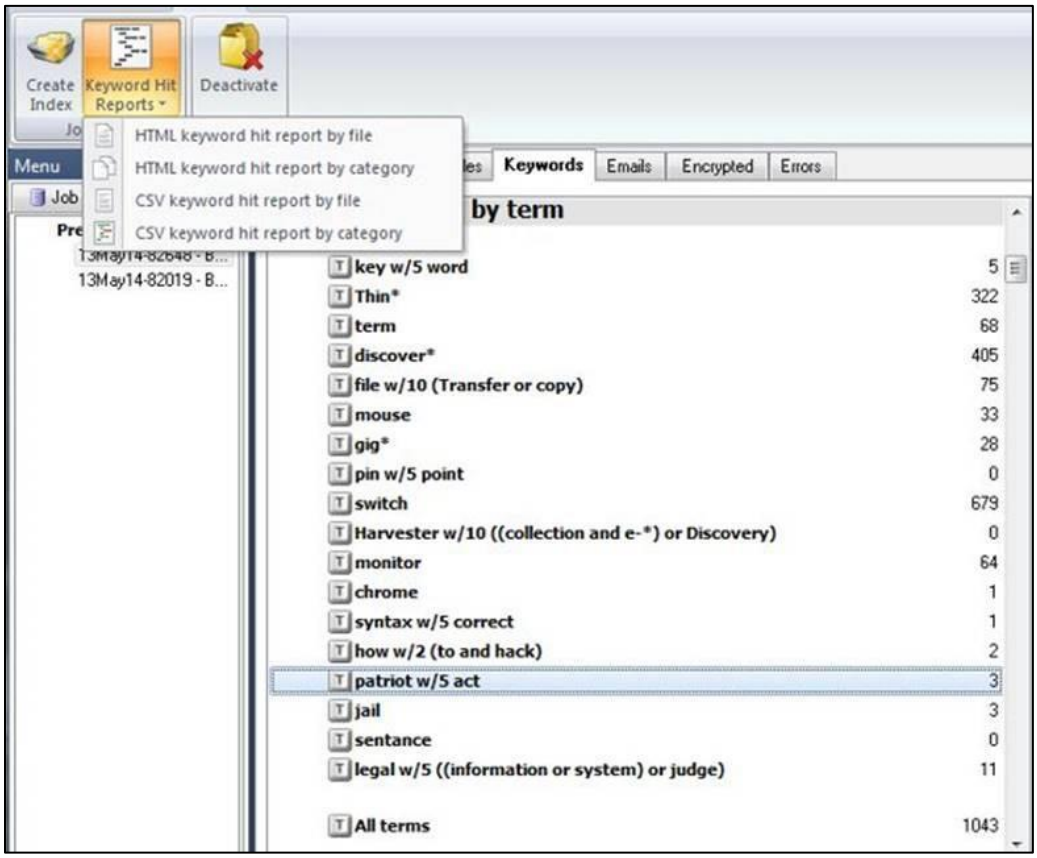
TAB NAME	DESCRIPTION
Summary	Contains run time statistics and totals for email and loose files categories.
Settings	A snapshot of the job profile settings. This can be very useful if users would like to know if, for example, they chose a setting or included all keywords.
Files	Tally for file types; includes total count and size.
Keywords	Lists total hits for each keyword entry and allows users to launch keyword hit preview.
Emails	Review which mail stores had matching items and the folder location.
Encrypted	Shows list of identified encrypted files organized by type.
Errors	Shows list of identified errors organized by category.

While the above descriptions should be self-explanatory, it is worth pointing out the extended functionality in the Keywords tab.

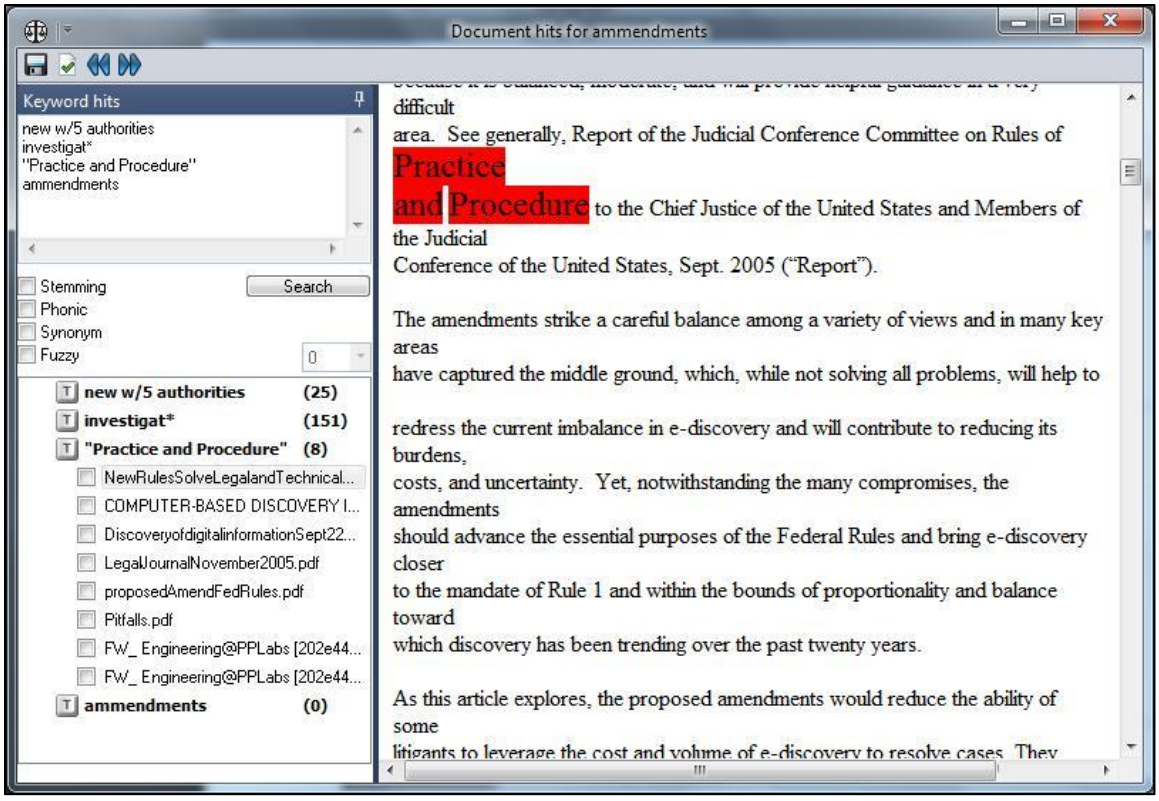
KEYWORDS AND HIT HIGHLIGHT REVIEW

Reference Video: ***Harvester Keyword Reports & Highlighting***

Harvester keyword hit reports and highlighted preview options are very useful for users who want to review search results. To take advantage of these powerful tools, users need to ensure an index was created for the job and keywords were chosen.



All entries will be listed in the **Keywords** tab as well at the total number of hits. Double clicking on an entry will bring up a window that lists the individual files with matching hits in the left pane. Clicking on a file in the list will display the contents in the preview windows and matching terms will be highlighted (as seen below).



The Keyword Hit Highlighter can automatically move to the next term that was found in the document by using the forward and backward arrows  at the top of the **Key Word Hit Highlighter** window.

KEYWORD PREVIEW FEATURES

The keyword preview window allows users to save keyword hit report by file list or category to an HTML or CSV file. To access these options click on the disk icon  in the upper left hand corner of the screen. After selection users will be allowed to browse to a location to store the file and provide a filename.

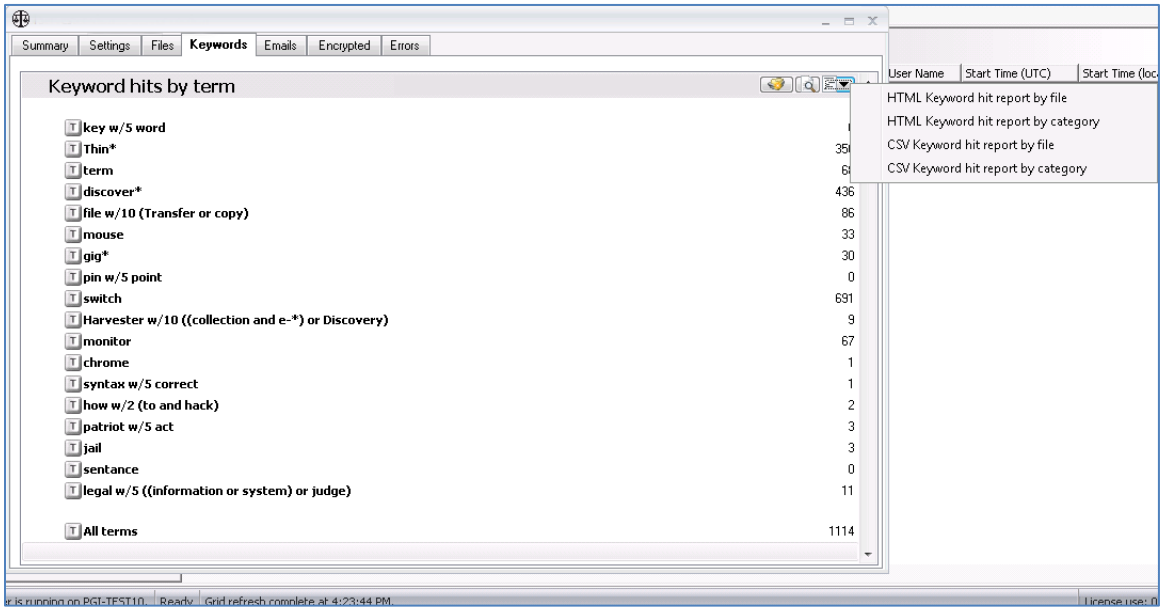
TAGGING FILES

The keyword list panel in the lower left hand corner of the keyword preview window enables users to:

- Click on an entry to preview the highlighted hits which are displayed in the right hand window
- Tag files that are of interest to the current review process by clicking the checkbox next to each item.
- Check multiple entries by using using CTRL+Shift (select multiple entries) **OR** holding Shift and clicking on 1st then last entry to select a range **OR** right-click and select ***Check All Items***.
- After selecting items right-click to ***Check selected items*** or ***Uncheck selected items***.
- Remove files from the keyword list by right-clicking and selecting ***Remove checked items*** or ***Remove un-checked items***. Users can also click on the check-mark icon in the upper left hand corner.
- Create file list from for items by right-clicking and selecting ***Create file list from checked items*** or ***Create file list from un-checked items***.

KEYWORD HIT REPORTING

In addition to previewing the hit results, users can create HTML or CSV reports of all hits, individual entries or selected items. To create hit reports for all items:



- Click on **Keyword Hit Reports**.
- Select the report you want to generate and the location where you would like to store the file.

Keyword hit reports by file will list all files and the location that match the entry.

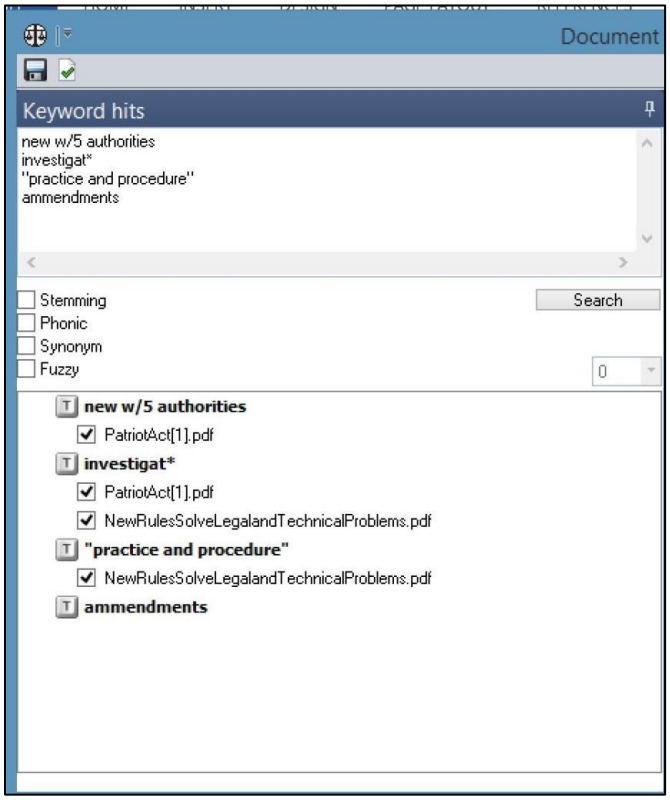
Keyword hit reports by category will create a tally report by file category.

To tag specific files relevant to your review and create a report, follow these steps:

- Double click on a specific term or **All Terms** to display a list of the selected documents. Using the following actions you can tag files:
    1. Click the check box next to each document.
    2. Shift or Control keys select individual or a range of hits and right click will mark or unmark the highlighted items.
- When finished tagging items click the  icon to remove the remaining items from the list.
- A report containing only the remaining items can be created by clicking the **Keyword Hit Reports** option in the toolbar of the document preview interface.

KEYWORD SEARCHING

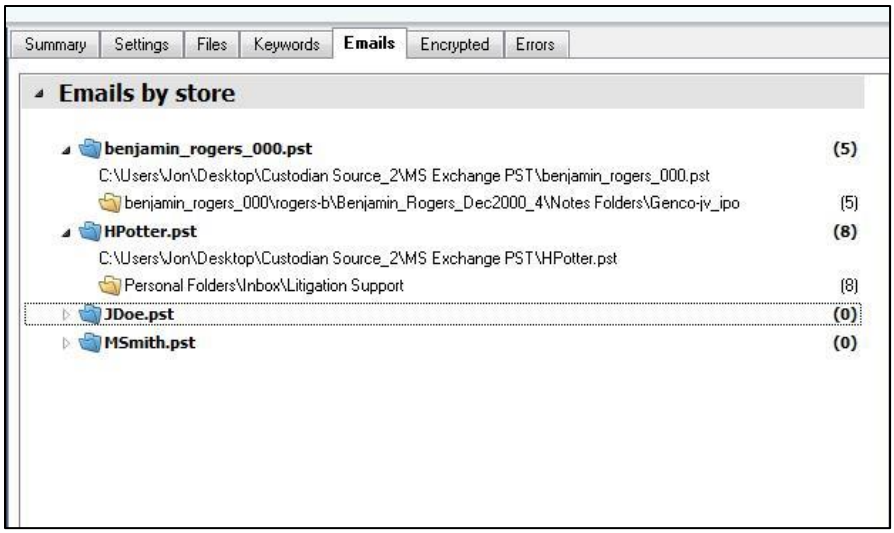
In addition to viewing the keyword results from a job, users can enter new search terms and review hits on-the-fly. This can be accomplished by entering the phrases in the upper left hand corner of the search hit preview screen. The keyword syntax format and rules are the same as available from the keyword tab in the job profile settings.



After clicking Search a new entry representing the phrase will appear as well as a list of the results. Clicking on an entry will load the contents in the preview window with the terms highlighted. Scrolling through the document may be required to see the hits.

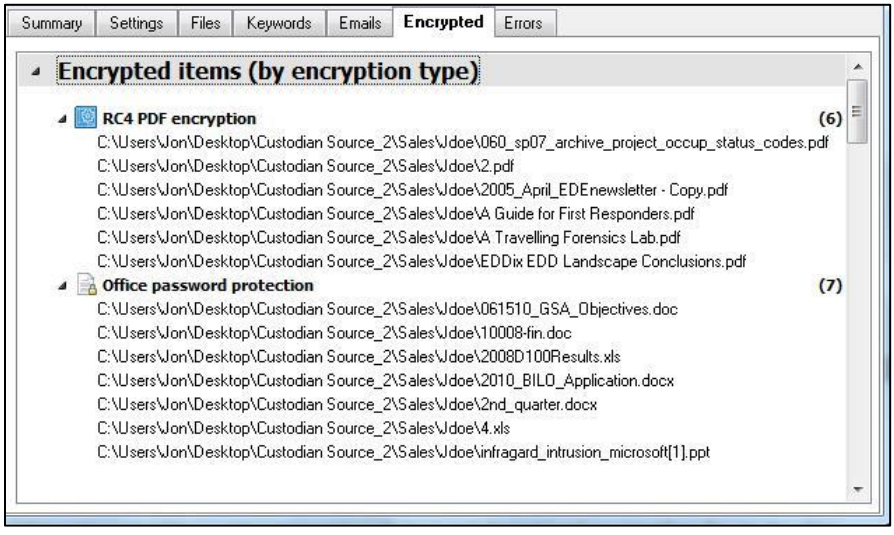
EMAIL RESULTS

If a user chooses to search emails or attachments within a mail store (PST, NSF) a list of the resulting matches by mail store will be displayed in the **Email** tab. Each mail store will be listed and allow users to expand to see the individual folders where the item are stored.



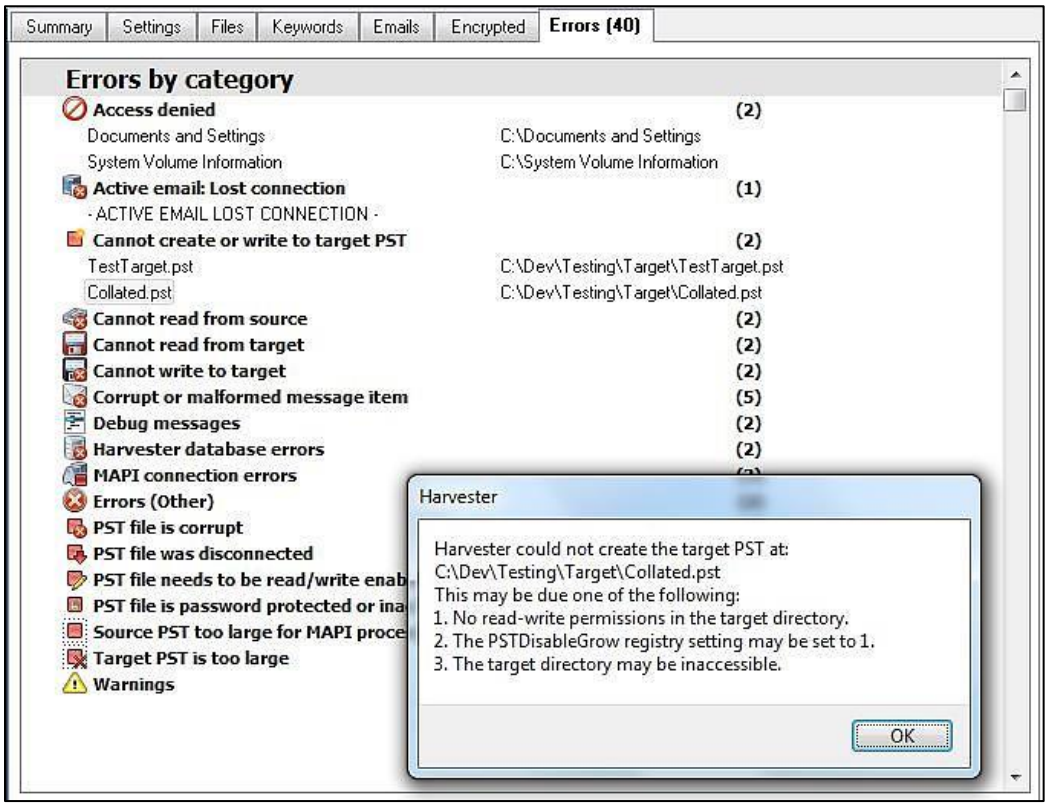
ENCRYPTED FILE RESULTS

If **Detect Encrypted Files** is selected, Harvester will check each file and tag those identified. The results will be displayed in the **Encrypted** tab. Each category can be expanded to see the individual file locations.



ERROR RESULTS

If errors occur during a Harvester job, they will be displayed in the **Errors** tab. Many common issues encountered by users are organized into categories and the total for each is displayed. Users can expand each category to see the individual file locations.



Double clicking on many of the individual items will display a message box (as seen above) that explains the error, common causes, and often how to fix the problem. The Harvester logs folder also contains a list of the errors encountered in ***_errors.log***.

SCRIPTING HARVESTER

Harvester includes functionality that allows users to create job files, start a Harvester job, and launch programs or utilities that can work with the data captured and files collected. The topics listed below cover how to control several aspects of Harvester.

HARVESTER VARIABLES

Several variables are available that can be used to automatically create directories and specify target computers. Here is a list of the current variables and how they are implemented.

Variable	Field(s)	Description
[SCDrive]	Log Path, Target Path, Common Log, Encrypted File Path, Collated Email Path	This variable is replaced with the drive letter that Harvester is running from, without the trailing slash. **Example:** ***[SCDrive]\Test1\Job*** would translate to: ***E:\Test1\Job*** if Harvester is running from somewhere on drive (E:)
[JobName]	Log Path, Target Path, Common Log, Encrypted File Path, Collated Email Path	This variable is replaced with the job name as defined in the Job Name field. **Example:** ***D:\Collections\[JobName]*** would translate to: ***D:\Collections\Brad Cowey Laptop*** if the job **Brad Cowey Laptop** is being run.
[CName]	Log Path, Target Path, Common Log, Encrypted File Path, Collated Email Path	This variable is replaced with the network name of the computer that the job is running on. **Example:** ***D:\Collections\[CName]*** would translate to: ***D:\Collections\MYLAPTOP*** if the job is running on a computer named ***MYLAPTOP***. This variable is very useful when running the same job from the same device on multiple computers.
[UName]	Log Path, Target Path, Common Log, Encrypted File Path, Collated Email Path	This variable is replaced with the username of the user that is logged in and running the software. **Example:** ***D:\Collections\[UName]*** would translate to: ***D:\Collections\JohnDoe*** if the user John Doe is logged into the computer and running the job. This variable is useful for separating data collected by many users accessing the same computer.
[Date]	Log Path, Target Path, Common Log, Encrypted File Path, Collated Email path	This variable is replaced with the current date (in local format). **Example:** ***D:\Collections\[Date]*** would translate to: ***D:\Collections\10-31-2010*** if the job is run on **October 31, 2010**.

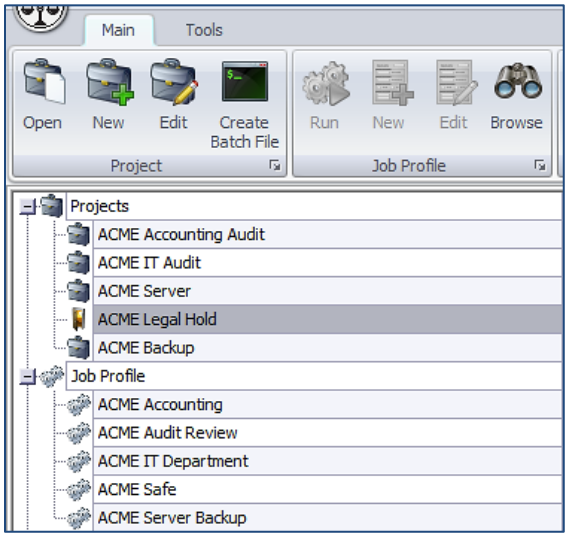
[DateTime]	Log Path, Target Path, Common Log, Encrypted File Path, Collated Email Path	This variable is replaced with the current date and time in the form dmmyy-tttttt. It is very useful for email-launched jobs since it will create a new set of folders for every run, even by the same user. **Example: *D:\Collections\[DateTime]*** would translate to: ***D:\Collections\4Apr11-135020*** if the job is run on **April 4, 2011 at 1:50:20 PM** local time.
[Logs]	Encrypted File Path, Collated Email Path	This variable is replaced with the path to the Logs directory after it has been translated at run time. This allows you to create subfolders to your logs folder for encrypted files or collated email stores.
[Target]	Encrypted File Path, Collated Email Path	This variable is replaced with the path to the Target directory after it has been translated at run time. This allows you to create subfolders to your target folder for encrypted files or collated email stores.
[LDrive]	Sources	This variable is replaced at run time with a new source for each logical drive connected to the computer running the job. The two drives that will not be added as sources are the drive that Harvester is running from and the drive that Harvester is copying files to. These exceptions are to prevent infinite loops.
[MDrive]	Sources	This variable is replaced at run time with a new source for each network folder that is mapped to a drive letter. For example, if the network folder ***\\NETSHARE1\Users\JohnDoe*** is mapped to drive letter **K:**, then drive **K:** is added as a source automatically at run time. A single instance of [MDrive] will add all network mapped drives as sources.
[UserFolder]	Sources	This variable is replaced at run time with the path to the My Documents (or equivalent user documents folder) for the logged in user only.
[UserFolders]	Sources	This variable is replaced at run time with the paths to all My Documents (or equivalent user documents folders) for all users on a machine. Note that you will not be able to copy files from these directories under all circumstances.
[PROMPT]	Sources	This variable will cause the program to prompt the user at runtime to drag and drop additional sources into the ESI Vault window.

LAUNCH JOB FILES FROM A COMMAND LINE

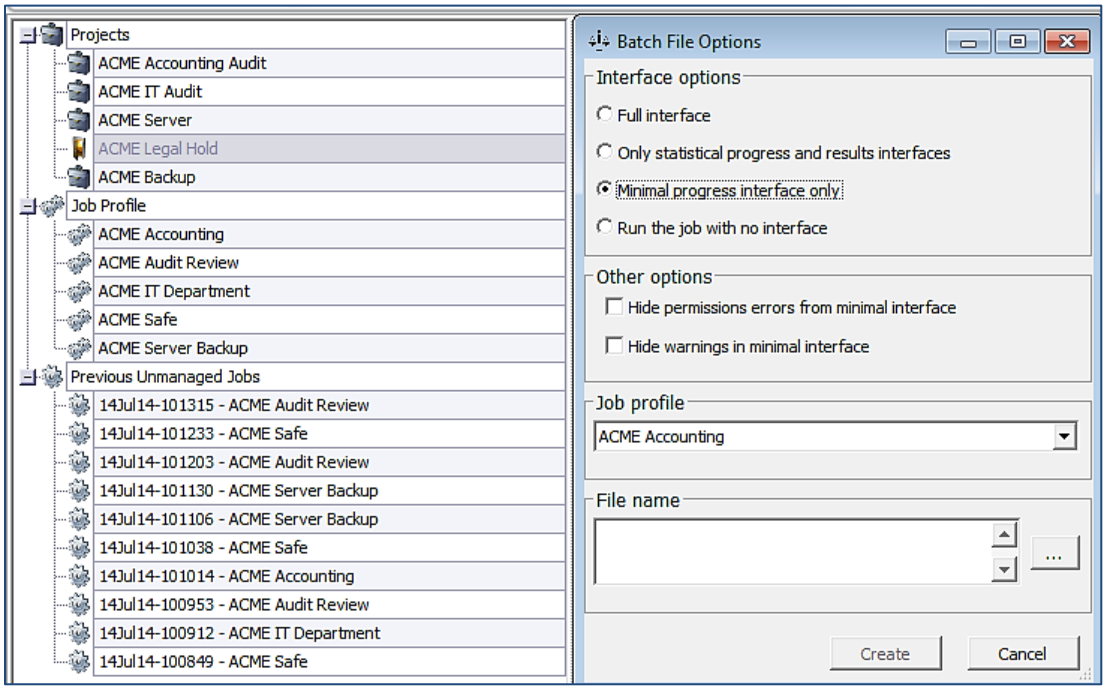
Harvester jobs can be automatically launched from the command line, batch files or applications that include *shell out* commands. Creating a batch file can be extremely useful in self-collection kits, legal holds, or streamlining everyday jobs. Harvester has many batch file creation options to fit your collection needs.

To create a batch file, follow the instructions below:

- Select the job you would like to create the batch file for.
- Once selected, click **Create Batch File** underneath the tools section of the toolbar.



- Select the options you would like your batch file to be created with. To see what each option's interface while running looks like, see **Batch File Interface Options** below.

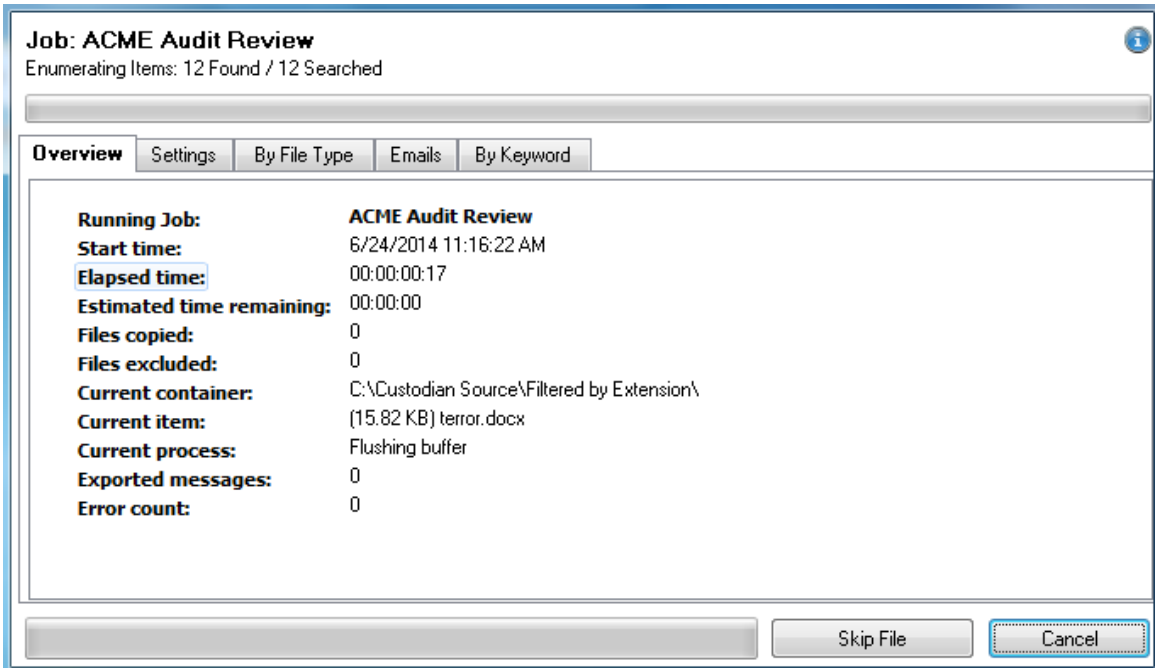


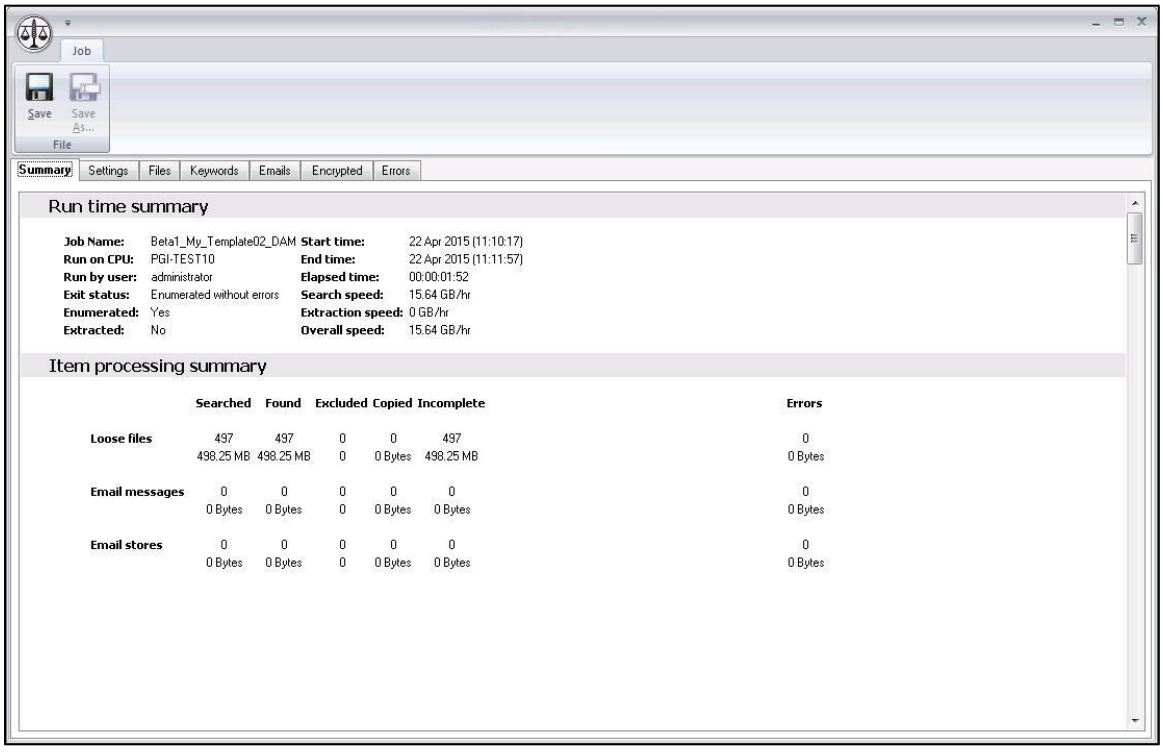
- When finished selecting the options for the batch file, click **Create** to finish the process. Your newly created batch file will be named ClickMe.bat unless otherwise named and will be located in the **Batch Files** folder in the location Harvester Server is running from.

BATCH FILE INTERFACE OPTIONS

Full Interface:

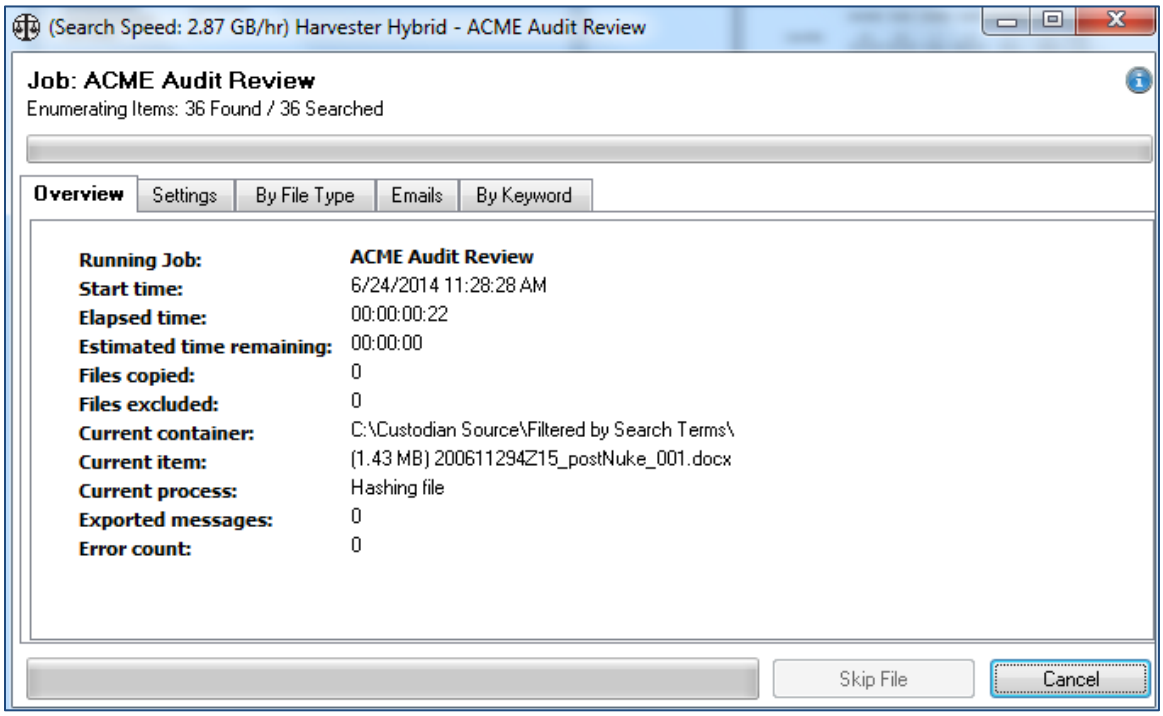
This option provides the full Harvester Interface when running and when viewing the history report when finished.

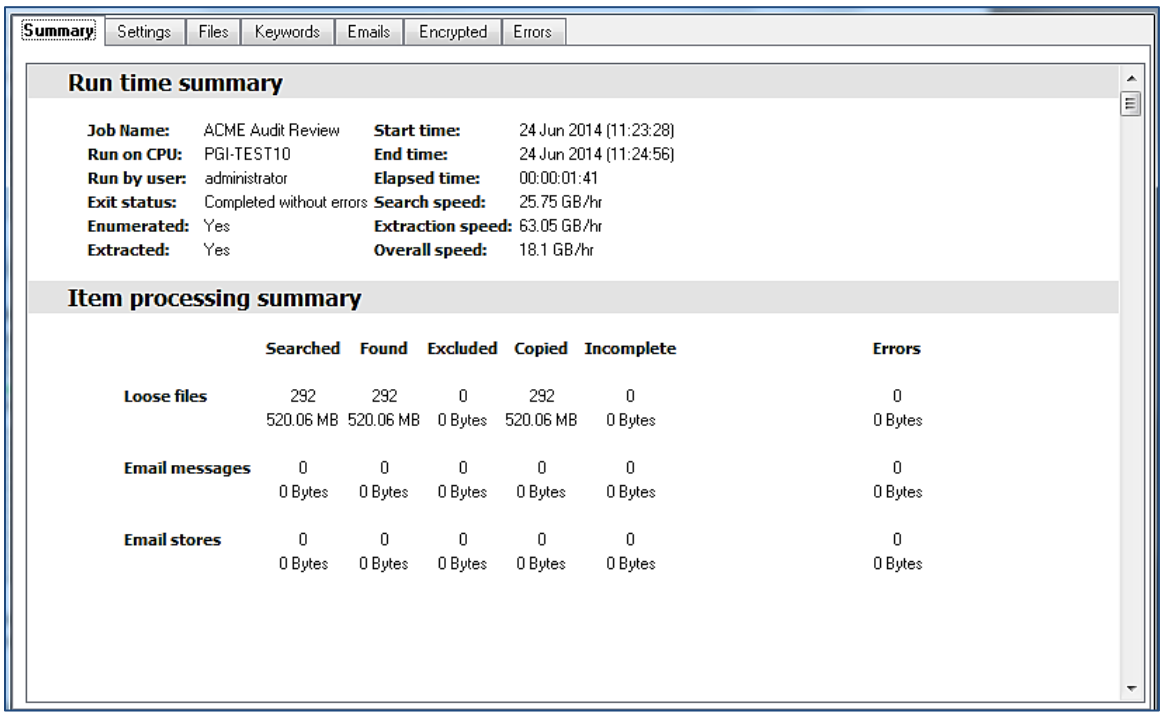




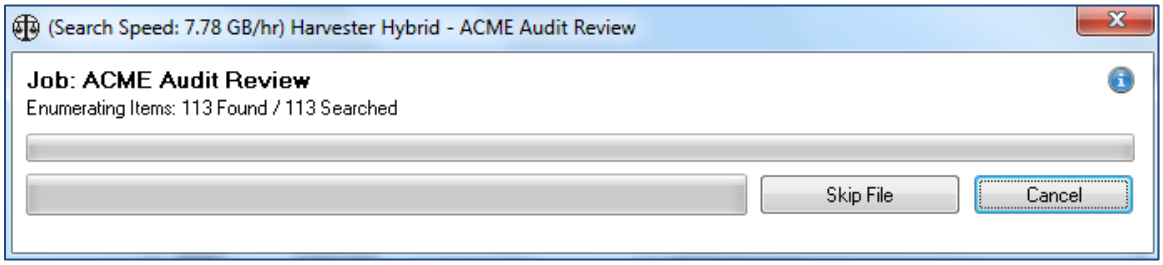
Only statistical progress and results interfaces:

This option provides only statistical progress and a limited history report when finished.

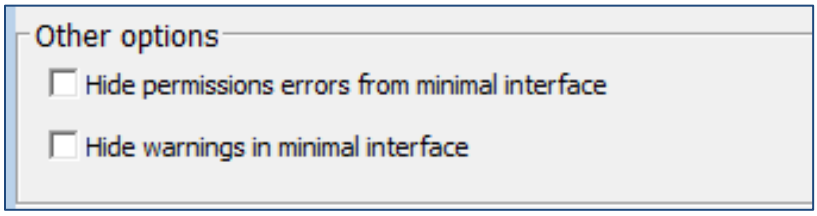




Minimal Progress Interface Only:
This option provides very little information about the running job and only provides a notice that the scan has finished, with no details about the job that was run.



NOTE: The *Other Options* only apply to the minimal progress interface setting.



Run this job with no user interface:
This option provides no user interface whatsoever. Once the batch file is double clicked, Harvester will run silently in the background with no indication that it is running.

When launching jobs directly from the command line, the available switches and required syntax for common scenarios are detailed below.

Harvester.exe [-as] [-q] [-silent] [-compact] [-suppress_permissions_alerts] [-suppress_warnings_alerts] [-stop] [-occ="path_to_occ_file"] [-resume="path_to_jobfile.scj"] [-retry]

-as	***autostart*** If there is only one file in the _occ directory in the application path, it will run it.
-q	***quiet*** This hides the job list window when a job is being started from the command line and instructs the application to quit once the job is completed. It is used in conjunction with either the -as, -occ, or -resume flags and has the effect of limiting the user interfaces to just the progress screen and the summary screen.
-silent	***Silent*** This hides all user interfaces. It is used in conjunction with the -as, -occ or -resume flags. This flag also forces the program to exit once the job is completed.
-compact	***Compact*** This hides all user interfaces except for the launch instructions, the ESI Vault, a basic progress bar, and an indication that the job has completed.

-suppress_ permissions_alerts	This flag, when used in conjunction with the ***-compact*** flag, will not treat permission errors as errors when alerting the user to errors at the end of the job. Permissions errors are still logged and visible in the history section and in the raw error log.
-suppress _warnings_alerts	This flag, when used in conjunction with the ***-compact flag***, will not treat non-critical warnings as errors when alerting the user to errors at the end of the job. Warnings are still logged and visible in the history section and in the raw error log.
-stop	**Stop** This closes the program after a job has been run.
-occ=	***Specify a job to run*** This allows you to specify the full path to an occ file to run. The path must be in quotes if it contains spaces, but may be in quotes even if spaces are not present in the path.
-resume=	***Resume a job*** This is used to resume a stalled job or to rerun errors on a job that has already been run (when used with -retry). The path to the _jobfile.scj must be specified. This will be located in the logs path of a job that has been started. If the path contains spaces, it must be in quotes. Quotes can also be used on paths that do not contain spaces.
-retry	***Retry errors*** This flag is used in conjunction with the -resume= flag and sets the error flags in the job database back to pending. This puts the job in a resumeable state where the errors are attempted again.

Command line scenarios

- While using Harvester, users want to launch a specific collection job from a batch file or other application. The Harvester job will immediately start and when completed bypass the job summary dialog box.

The ***-silent*** switch is used to automatically start a specific job and suppress all message and progress dialog boxes. This switch also tells Harvester to bypass the job summary message when the job is completed. The following syntax is used: *[Harvester executable path/filename] [switch] -occ= [occ job file path OR occ job name].*

Example:
"\\MyServer\Harvester\Harvester.exe" -silent -occ="\\MyServer\Harvester _occ\profile1.occ"
or
"\\MyServer\Harvester \Harvester.exe" -silent -occ="profile1.occ".

The second option only requires the name of the job file. When processing this command, Harvester will check the _occ directory located in the path already specified.

- ***To run a specific job with minimal user interface that stops after the job has been run:***
  D:\Harvester\Harvester.exe -q -occ="D:\Harvester_occ\SampleJob.occ"
- ***Or, if the job file is on a network share:***
  D:\Harvester\Harvester.exe -q -occ="\\FileServer34\Legal\HarvesterJobs\SampleJob.occ"
- ***To rerun errors on a job and stop after it's finished:***
  D:\Harvester\Harvester.exe -q -resume="D:\Logs\12June2012_jobfile.scj" -retry
- ***To run a job with no user interface:***
  D:\Harvester\Harvester.exe -silent -occ="D:\Harvester_occ\SampleJob.occ"

- ***To rerun errors with no user interface:***
  D:\Harvester\Harvester.exe -silent -resume="D:\Logs\12June2012_jobfile.scj" -retry
- ***To launch a specific job with minimal interfaces that does not report permissions errors to the user:***
  D:\Harvester\Harvester.exe -compact -occ="SampleJob.occ" -suppress_permissions_alerts

The first scenario is based on automatically launching and running Harvester in a 'stealth' mode that is completely automatic and doesn't require any further user interaction. This will most commonly be used to automate network collections. It is important to review any error logs to make sure that the job completed successfully, and no further action is required.

LAUNCHING THIRD PARTY UTILITIES USING "SHELL OUT" COMMAND

Using the ***Shell Command to Execute on Job Start***, or ***Shell Command to Execute on Job Completion*** options to launch other applications or utilities allows you to automate processing jobs. The shell out command will be executed when a job is launched or completed (respectively).

The following steps specify how to add a shell out command to an .occ job file:

- Launch Harvester
- Open the .occ job profile
- Click on ***Advanced*** button at the bottom of the job manager
- Enter the commands in the text boxes provided and click ***OK***

Using the Command to run at job start, or Command to run at job end options to launch other applications or utilities allows you to automate processing jobs. The command will be executed when a job is launched or completed, respectively. With Show command prompt window selected, the Windows Command Prompt window will open when the command is executed.

An example of a command to run would be ***C:\Windows\System32\notepad.exe***, when entered into the Command to run at job end will open **Notepad** when the Harvester job has finished.

Another example would be entering ***C:\Program Files (x86)\Microsoft Office\Office14\outlook.exe*** in the Command to run at job start, which will open **Microsoft Office Outlook** when the Harvester job is launched.

Logs Files

The following is a list of Harvester log files and what they contain. These log files (if relevant to the job) will appear in the log folder that was specified under the Targets tab in the Harvester Job Profiles interface.

_jobfile.scj	Contains job settings and a list of the sources processed. This file is used to resume jobs**.**
_job.sdb	Is a data file which contains information captured during enumeration and copying
_flist_errs.log	Errors associated with a File list used as a Source
ComputerInfo.txt	Text log containing information about the computer running Harvester, such as computer name, user name, operating system, processors, and attached drives.
_verification_log.csv	Contains a list of files copied, associated metadata, and any errors encountered. The verification log acts as the ***chain of custody*** for loose files.
_email_verfication_log.csv	Contains a list of emails copied, associated metadata, and any errors encountered. The email verification log acts as the chain of custody for emails.
_errors.log	Contains a list of errors encountered. Although these are in the verification log, a separate file is created so users can easily review just the errors and use the log to reprocess files.
exclusions.log	Contains a list of files which were excluded as a result of DeNISTing, de-duping, and file type filtering as well as the reason for the exclusion.
_ts_mismatch.log	There can be slight discrepancies (fractions of a second, or possibly a few seconds) in the file system timestamps on the copied files when the file systems on the source are different than the destination. File systems store the time in different *resolutions* so an exact match may not be possible. Discrepancies are common when copying from a file system with high timestamp resolution (NTFS) to one with a lower timestamp resolution (FAT32).Since an *error* message will be logged and displayed for each file, a separate file is created to store the messages, so the primary error log is used to store messages related to incomplete copies.
_silent.log	Created when Microsoft Windows error occur which are not related to a specific file. This log may also contain notification or warning information for other types of errors
suspect.log	Includes a list of files where the header signature doesn't match the expected extension. This log may be created when the ***File Types*** categories are selected which rely on file header signatures.
filelist.txt	Created when the ***Create File List*** option is enabled. The filelist.txt contains a list of responsive files from the data sources selected. The log contains one file path per line.
folderlist.txt	Created when the ***Create Folder List*** option is enabled. The folderlist.txt contains a list of folders from defined data sources. The log contains one folder path per line.
tally.txt	The tally.txt file contains the total number of files and size for the selected data sources as well as general statistical information about a job that has been completed or cancelled.
_extension_tally.csv	Created when the ***Create Tally Summary*** option is enabled. The _extension_tally.csv file contains a statistical breakdown of each file extension encountered by count, by size, and by whether it was a loose file, in an archive, or attached to an email.
email_attachment_list_extended.txt	This is a tab separated text file that contains the following values for each email attachment found: Path to PST>>InternalPath/subject of email, name of attachment, date created, date received (Date created and date sent will be the same for non-Exchange attachments)
file_list_extended.txt	This is a tab separated text file that contains the following values for each logical file that was found: Full path of the file, file name, date created, date last modified
email_list_extended.txt	This is a tab separated text file that contains the following values for each email that was found: Path to PST>>Internal Path/subject of email, date sent, date received
_email_attachments.csv	This is a comma separated values file containing columns for the following information: Path to the PST, Internal path to the message, subject of the message, attachment name, and attachment size.
_encrypted_files.txt	This is a text file that contains the paths of all of the files that were deemed encrypted, image-only or unsearchable. There is one path per line.
_encrypted_email_attachments.csv	This is a comma separated values file that contains the path, subject, container info, and attachment information for any email attachments determined to be encrypted. This log is produced when encryption detection is enabled and email attachments are being searched.
_image_only_pdfs.log	This is a list of file paths for pdf files that were determined to contain only image data but

	are not otherwise encrypted. This log is produced when key word searching and encrypted file detection are both employed.
_duplicate_emails.log	This is a text file that contains the email path and subject, as well as the original PST it was located in for any duplicate emails that have been found. This log is only produced when the ***Exclude duplicates*** email option has been checked.
hashlist.md5	This is the sorted MD5 hash list that is produced when the ***Create hash list*** option is checked. It contains only hashes for loose files, not for emails or attachments.
emails_hashlist.md5	This is the sorted MD5 hash list that is produced when the Create hash list option is checked and emails are being searched. It contains only hashes for emails that were responsive.
longpaths_source.log	This is a text file that contains any source paths that are greater than 255 characters.
longpaths_dest.log	This is a text file that contains any destination paths that are greater than 255 characters. These are logged because these files may be difficult to get to via normal means.
_nonsearchable_email_attachments.csv	This is a comma separated values file that lists any email attachments that came up as non-searchable during a key word search of email attachments. The file lists the following properties of each attachment: The path to the email store it was found in; The entry ID of its parent message; The folder within the store where the message can be found; The subject line of the parent email; and the file name of the attachment.
_job.sdb (PPLM)	contains data for emails encountered in a particular email store
_jobfile.scj (PPLM)	contains the settings and instructions for a particular email store
calling_command.txt (PPLM)	contains the instructions used to launch a particular email store
Enum_Exit.txt (PPLM)	contains the exit conditions at the end of enumeration for a particular email store
Error.txt (PPLM)	currently not used, but it will contain error information not related to debug functions within the processing of a particular email store
Log.txt (PPLM)	contains setup, version, and logging information for the enumeration and collection of a particular email store
MailComplete.txt (PPLM)	contains exit conditions for the last email thread process for a particular email store
Progress.txt (PPLM)	contains the last progress message for the thread processing of a particular email store

Understanding OCC Files

.OCC/.SCJ FILE STRUCTURE DEFINITION

To open an existing job file, browse to the _occ sub-directory located in the Harvester application directory (where files were unzipped). Double click or right click on one of the files and select notepad or your preferred text editor to open. You should see something similar to the content below:

[JNAME] Brad Cowley Laptop
[JINSTRUCT] Email, Microsoft Office Documents, & PDF files from serial# 5J8I9RT67
[JERRINST] If you have any questions regarding this project please contact Leroy Jenkins 402.555.1212
[EDAPAUSE] 0
[THREADS] 0
[SUSPECTNOCOPY] 0
[WRITE_TO_VHD] 1
[VHD_PATH] C:\VHD\MyVHD.vhd
[VHD_TARGET] _Target
[VHD_LOGS] [SCDrive]_Logs

[VHD_MOUNTPOINT] C:\Users\Admin\AppData\Local\Temp\OCC2976.tmpmnt
[TARGET] [SCDrive]\[JobName]\
[JPATH] [SCDrive]_Logs\[JobName]\
[SRC] C:\
[ENGINE] SC
[SEARCHCREATED] 0
[SEARCHMODIFIED] 0
[SEARCHACCESSED] 0
[SILENT] 0
[OVERWRITE] 0
[COPYFILES] 1
[TALLYSUM] 1
[FILELIST] 1
[FOLDERLIST] 1
[RENAME] 0
[FULLPATH] 0
[ROOTFOLDERS] 0
[COPYEMPTIES] 1
[COPYSUBS] 1
[ZIPDIRS] 1
[SEPARATE_TS] 0
[SUSPECTNOCOPY] 0
[EXTOP] 0
[LOGEXCLUSIONS] 1
[LOG] 1
[S-HASH] 1
[D-HASH] 1
[STARTEX]
[SHELLEX]
[SEARCHEMAILS] 0
[KWSUBJECTBODY] 0
[ACTIVEEMAIL] 0
[KWATTACH] 0
[KWSTEMMING] 0
[KWPHONIC] 0
[KWSYNONYM] 0
[KWFUZZY] 0
[KWARCHIVEOPT] 0
[HASHFILTER] 0
[HFILTER_ATTACH] 0
[HASHFILTERINCLUDE] 0
[HASHFILTEREXCLUDE] 1
[SKIPSYSFILES] 0
[SKIPSYSDIRS] 0
[SKIPTEMP] 0
[KWGO] 0
[KWONLY] 1
[KWHITENCRYPTED] 0
[DETECTENC] 0
[COPYENCTO] 0
[ENCTARGET] [SCDrive]\[JobName]\encrypted
[ENCFULLPATHS] 0
[ENCROOTFOLDERS] 0

[ENCSUBFOLDERS] 1
[COPYENCNORMAL] 0
[SKIPDUPES] 0
[HASHLIST] 1
[EDEDUPE] 0
[EATTACHDATES] 0
[EADDYEX] 0
[LNADDYEX] 0
[EPTYPE] 0
[EEXPORTFORMAT] 3

A description of the .occ /.scj fields is listed below:

In fields that require a 1 or 0 **1*=True/Checked*** and ***0=False/Unchecked***

FIELD	VALUES	NOTES
[PROD]	text	(Job File Only, Automatic) The product name that generated the job file
[VERSION]	text	(Job File Only, Automatic) The version number of the product that generated the job file
[APPPATH]	text	(Job File Only, Automatic) The path to the executable that generated the job file
[EXEC_CPU]	text	(Job File Only, Automatic) The executing machine name.
[EXEC_USER]	text	(Job File Only, Automatic) The executing user name.
[JNAME]	text	(Required) This is the job name. For best results, keep simple and only use values that can be used in a file path.
[JINSTRUCT]	text	(Optional) Contains job description (up to 255 characters) and is displayed in the job list and startup message box. This only appears in OCC files and there will be one [JINSTRUCT] entry per line of information to be displayed. This only appears in OCC files and there will be one [JERRINST] entry per line of information to be displayed.
[JERRINST]	text	(Optional) Contains contact information for project manager and/or procedures to follow in case of errors (up to 255 characters). Information is displayed in the startup message box and at the end of the job if errors are encountered.
[RUNNING_AS_ADMIN]	text	(Job File Only, Automatic) *True* or *False* Indicates whether or not administrator credentials were used to launch the job.
[DATA_ASSESSMENT_MODE]	text	(Job File Only, Automatic) *On* or *Off* Indicates whether data assessment mode is on or off for the instance of the job that produced this job file. Resuming a job with this value set to On will change it to Off.
[WRITE_TO_VHD]	1 or 0	(Automatic) Controls whether a VHD file container is used as a target.
[VHD_PATH]	text	(Optional) Contains the VHD container file path.
[VHD_TARGET]	text	(Optional) Contains the target path within the VHD file.
[VHD_LOGS]	text	(Optional) Contains the logs path used with the VHD options.
[VHD_MOUNTPOINT]	text	(Job File Only, Automatic) Contains the VHD container file mount point path.
[TARGET]	text	(Required) By default contains variables that will create a subdirectory, using the [JNAME] data, on the root of the drive where Harvester is running. The collected files are copied to this [TARGET] directory. Other variables, a network path (UNC) or hard path can be used. In the SCJ file, any variables are translated to their run time values.
[JPATH]	text	(Required) In the OCC file, this field contains variables and path information that will create a *_Logs* directory. Logs are stored in this directory. Other variables, a network path (UNC) or hard path can be used. In the SCJ file, any variables are translated to their run time values
[EDAPAUSE]	1 or 0	(OCC File Only, Optional) This indicates whether the job is to run in Data Assessment Mode, enumerating items and pausing for statistical or other reports before being resumed for the copy phase. 1=Pause after enumeration. 0=Continue to copy phase after enumeration.
[SHADOW]	text	(Job File Only, Automatic) Contains the temporary shadow volume information (Volume Shadowed, mount location, and GUID) for a single shadowed volume.
[VSS_PRESENT]	text	(Job File Only, Automatic) *True* or *False* Indicates whether VSSADMIN.exe is present.
[VSS_AUTHORIZED]	text	(Job File Only, Automatic) *True* or *False* Indicates whether admin credentials were used to

		launch the job.
[SRC]	text	(Required) Contains one data source which can consist of drive letters, directories and files or file list. A single job file may have many [SRC] entries. In an OCC file, this may contain variables such as [LDrive]. In the SCJ file produced when the job is run, these variables are translated to their run time values and may produce additional [SRC] entries.
[MSRC]	text	(Job file only) This denotes a source that was added manually at run time by the user, using the ESI Vault.
[PROMPT]	1 or 0	(OCC file only) This field indicates whether or not the ESI Vault will appear at run time to allow users to add additional sources. 1 = ESI Vault will appear. 0 = The job will run with no ESI Vault window.
[HAS_OUTLOOK]	text	(Job File Only, Automatic) *True* or *False* Indicates whether the computer running the job had MAPI-enabled, 32-bit Outlook installed.
[HAS_LOTUS]	text	(Job File Only, Automatic) *True* or *False* Indicates whether the computer running the job had Lotus Notes installed
[ENGINE]	SC	(Required) Must be *SC*
[EXCLUSION]	text	(Optional) Contains a single path-based exclusion pattern. One job file may have multiple [EXCLUSION] entries.
[FNAMEFILTER]	text	(Optional) Contains a single path-based inclusion pattern. One job file may have multiple [FNAMEFILTER] entries.
[COPYEMPTIES]	1 or 0	(Required) Controls whether empty sub directories are copied.
[COPYSUBS]	1 or 0	(Required) Controls whether subdirectories under the selected data source are copied.
[COPYFILES]	1 or 0	(Required) Controls whether files are copied. This will be set to 0 if the user wants to generate a 'file list' or 'tally' report without copying files. There is no interface to set this value, but setting it by changing the value in the OCC file will prevent the files from being copied, even when resuming a job. This is equivalent to Data Assessment Mode without an option to continue after enumeration.
[TALLYSUM]	1 or 0	(Required) Controls whether a job summary report is generated.
[FILELIST]	1 or 0	(Required) Controls whether a file list report is created.
[FOLDERLIST]	1 or 0	(Optional) Controls whether a folder list will be created.
[BMONTH]	NUM	(Optional) Beginning month range for MAC time filtering
[BDAY]	NUM	(Optional) Beginning day range for MAC time filtering
[BYEAR]	NUM	(Optional) Beginning year range for MAC time filtering
[EMONTH]	NUM	(Optional) Ending month range for MAC time filtering
[EDAY]	NUM	(Optional) Ending day range for MAC time filtering
[EYEAR]	NUM	(Optional) Ending year range for MAC time filtering
[SEARCHCREATED]	1 or 0	(Required) Controls whether Date Created is used for a date search.
[SEARCHMODIFIED]	1 or 0	(Required) Controls whether Date Modified is used for a date search.
[SEARCHACCESSED]	1 or 0	(Required) Controls whether Date Last Accessed is used for a date search.
[SILENT]	1 or 0	(Required) Controls whether windows errors are (1) logged to a separate file or (0) shown in a popup box
[OVERWRITE]	1 or 0	(Required) Controls whether the Overwrite option is selected in the file collision options.
[RENAME]	1 or 0	(Required) Controls whether the Rename option is selected in the file collision options.
[FULLPATH]	1 or 0	Controls whether or not the target paths will (1) reflect the full source paths above their original root directories or (0) reflect only the folders below the folder defined in the source.
[ROOTFOLDERS]	1 or 0	(Required) Controls whether root folders (drive letters) are included in job path.
[COPYEMPTIES]	1 or 0	(Required) Indicates whether the *Copy Empty Folders* box was checked. A value of 1 indicates that folders in the source that contained no hits will be represented in the target. A value of 0 indicates that they will be left out.
[ZIPDIRS]	1 or 0	(Required) Indicates whether the *Process Zip files as directories* box was checked. If this value is set to 1, then the contents of zip files will be subject to the defined file filters.
[SUSPECTNOCOPY]	1 or 0	(Required) Indicates whether the *Do not copy files with suspect extensions* box was checked. If this value is 1, then files whose extensions do not match their headers will be logged, but will not be copied. This is only applicable when using header/file type filtering.
[CREATESUBS]	1 or 0	(Required) Controls whether subdirectories in the job path are created. 1 indicates that they will be created. 0 indicates that all responsive files will go into the same target folder.
[SEPARATE_TS]	1 or 0	(Optional) Legacy option. 1 indicates A separate log for time stamp discrepancies is created automatically. 0 indicates that time stamp discrepancies will be considered copy errors.
[EXTS]	text	(Optional) Contains the specifications listed in the file type/extensions box
[EXTLIST]	text	(Optional) Contains the path to the text file containing a list of file extensions to use for

		processing.
[EXTOP]	1 or 0	(Required) 0=Include specified extensions/types. 1=Exclude them.
[LOG]	1 or 0	(Required) Controls whether a verification log is created. 1 indicates that the verification log will be created. 0 indicates that the verification log will not be created.
[LOGEXCLUSIONS]	1 or 0	(Required) Controls whether an exclusion log is created. 1 indicates that an exclusion log will be created. 0 indicates that an exclusion log will not be created.
[S-HASH]	1 or 0	(Required) Controls whether the source file is hashed for verification. 1 indicates that all source file hashes will appear in the verification log. 0 indicates that the source file hashes will not be listed in the verification log.
[D-HASH]	1 or 0	(Required) Controls whether the destination file is hashed for verification. 1 indicates that all destination file hashes will appear in the verification log. 0 indicates that the destination file hashes will not be listed in the verification log.
[STARTEX]	text	(Optional) Contains a shell command to run at the beginning of the job.
[SHELLEX]	text	(Optional) Contains a shell command to run at the end of the job.
[SHOWSHELL]	1 or 0	(Optional) Controls whether or not a command line window will be opened to run the job start and job completion commands. The default is '0' – No window.
[SEARCHEMAILS]	1 or 0	(Required) Controls whether PST files that are encountered in the search should be searched as email containers. 1 indicates that loose PST files should be searched as email containers. 0 indicates that they should be treated as normal loose files.
[SEARCHLIVEPST]	1 or 0	(Required) Controls whether PST files that are mounted in the default Outlook profile should be searched. 1 = yes. 0 = no.
[SEARCHEXCHANGEBOX]	1 or 0	(Required) Controls whether any Exchange Mail Boxes connected to the default Outlook profile should be searched. 1 = yes. 0 = no.
[SEARCHPUBLICFOLDERS]	1 or 0	(Required) Controls whether any Exchange Public Folders connected to the default Outlook profile should be searched. 1 = yes. 0 = no.
[SEARCHLOTUS]	1 or 0	(Required) Controls whether NSF files that are encountered in the search should be searched as email containers. 1 = yes. 0 = no.
[ACTIVELOTUS]	1 or 0	(Required) Controls whether the default mail store that the current user connects to via Lotus Notes should be searched. 1 = yes. 0 = no.
[DRIVETOUNC]	1 or 0	(Optional) Indicates whether the *Translate mapped network drives to UNC* box has been checked. 1 = Any mapped drive letters that attach to UNC paths will be translated to those UNC paths. 0 = The mapped drive letters will be used.
[FILTERESIV]	1 or 0	(Optional) Indicates whether the *Apply filters to user-added folders* box is checked. 1 = Filters will be applied to folders that were dragged and dropped into the ESI Vault. 0 = Folders that were dragged and dropped into the ESI Vault will be copied verbatim without applying filters.
[HAS_OUTLOOK]	1 or 0	(Optional, no interface) Marks if the source included an Outlook email file.
[HAS_LOTUS]	1 or 0	(Optional, no interface) Marks if the source included a Lotus Notes email file.
[KWSUBJECTBODY]	1 or 0	(Optional) Indicates whether the *Use Key Word Filter for email subject/bod*y box was checked. *1 = Email subjects and bodies will be searched using the defined keyword filters. 0 = Email subjects and bodies WILL NOT be searched using the defined keyword filters*
[KWEMAILHEADERS]	1 or 0	(Optional) Indicates whether the Search Email Headers box is checked. 1 = Email headers will be searched using the defined keyword filters. 0 = Email headers WILL NOT be searched using the defined keyword filters.
[KWCREATEINDEX]	1 or 0	(Optional) Indicates whether the *Create Index* box is checked. 1 = A keyword index will be created. 0 = No keyword index will be created.
[CACHEINDEX]	1 or 0	(Optional) Indicates whether the *Create Index box* is checked. 1 = File contents will be cached in the index. 0 = No file contents will be cached in the index.
[KWATTACH]	1 or 0	(Required) Controls whether email attachments will be subject to the defined keyword filters. 1 = Yes. 0 = No.
[KWEXCLUDE]	0 or 1	(Required) Controls whether or not a key word hit triggers an exclusion of the item from the list of responsive items. *0* indicates that the item will be included. *1* indicates that the item will be excluded.
[KWSTEMMING]	1 or 0	(Required) Controls whether stemming should be used in key word searching. 1 = Yes. 0 = No.
[KWPHONIC]	1 or 0	(Required) Controls whether phonic matches should be included in key word searches. 1 = Yes. 0 = No.
[KWSYNONYM]	1 or 0	(Required) Controls whether synonym matches should be included in keyword searches. 1 = Yes. 0 = No.
[KWFUZZY]	1 or 0	(Required) Controls whether fuzzy matches (misspellings) should be included in key word searches. 1 = Yes. 0 = No.
[KWFUZZYTOL]	1-10	(Optional – Only required if [KWFUZZY] is *1*) Controls which value is selected for fuzzy tolerance (how misspelled a word is)
[KWARCHIVEOPT]	NUM	(Required) This value reflects which option is selected for handling keyword hits within an

		archive file. 0=Copy whole archive on match. 1=Extract matching files
[HASHFILTER**]**	1 or 0	(Optional) Controls if ***hash filter*** option is selected in Harvester
[HASHFILTERINCLUDE]	1 or 0	(Required if [HASHFILTER] = 1) Controls whether only files with listed hashes will be included in the results. If both this value and the [HASHFILTEREXCLUDE] value are set to 1, then files with listed hashes will be excluded.
[HASHFILTEREXCLUDE]	1 or 0	(Required if [HASHFILTER] = 1) Controls whether only files without listed hashes will be included in the results. If both this value and the [HASHFILTERINCLUDE] value are set to 1, then files with listed hashes will be excluded.
[KWGO]	1 or 0	(Required) Indicates whether loose files will be subject to keyword search filters. 1 = Yes. 0 = No.
[KWONLY]	1 or 0	(Required) A value of 1 indicates that any files that are not key word searchable should not be included in the results, except for defined exceptions.
[KWHITENCRYPTED]	1 or 0	(Optional) Indicates whether the *Count Encrypted and Image-only files as KW hits* box was checked. Not necessary if [DETECTENC] is *0. 1* = Encrypted items are counted as hits. *0 = Encrypted items are not counted as hits.*
[KWEXCEPTIONS]	text	A comma-separated list of file extensions that should be included even though they are not key word searchable. This setting only applies if the [KWONLY] value is *1*.
[SKIPSYSFILES]	1 or 0	(Required) Controls if system files will be skipped. 1 = Files with the system attribute set will be excluded. 0 = The system attribute flag will not be evaluated.
[SKIPSYSDIRS]	1 or 0	(Required) Controls if system directories will be skipped. 1 = Directories with the system attribute set will be excluded. 0 = The system attribute flag will not be evaluated for directories.
[SKIPTEMP]	1 or 0	(Required) Controls if ~~system~~ temporary files will be skipped. 1 = Files with the temporary attribute set will be excluded. 0 = The temporary attribute flag will not be evaluated.
[SKIPDUPES]	1 or 0	(Required) Controls if duplicate loose files are excluded. 1 = Duplicate files are logged, but excluded. 0 = The duplicate status of files will not be evaluated.
[HASHLIST]	1 or 0	(Required) Controls if hash lists will be used for filtering. 1 = Hash lists will be loaded and each file will be hashed for comparison. 0 = No hash lists will be loaded.
[HFILTER_ATTACH]	1 or 0	(Optional) Indicates whether the *Apply to Email Attachments* box was checked. Not necessary if [HASHFILTER] is *0* or if emails are not being searched. 1 = Hash list filtering will apply to email attachments. 0 = Hash list filtering WILL NOT apply to email attachments.
[KWLIST]	text	(Optional) Contains a single keyword filter entry (term). A single job file may have many [KWLIST] entries.
[EADDY]	text	(Optional) One or multiple entries that contain each line in the Address/Domain to Search For section of the Loose Outlook PST filtering.
[AEADDY]	text	(Optional) One or multiple entries that contain each line in the Address/Domain to Search For section of the loose PST search filter options.
[LNADDY]	text	(Optional) One or multiple entries that contain each line in the Address/Domain to Search For section of the Lotus Notes and Active Lotus search filter options.
[PSTSRCHFOLDER]	text	(Optional) One or multiple entries that contain filters identifying which PST folders to search when searching loose PST files.
[AESRCHFOLDER]	text	(Optional) One or multiple entries that contain filters identifying which email folders to search. When searching Exchange mailboxes, Exchange public folders, or mounted PST files.
[EDEDUPE]	1 or 0	(Required) Controls whether email de-duping is enabled for emails encountered in loose PST files.
[AEDEDUPE]	1 or 0	(Required) Controls whether active Outlook email (Exchange, Public Folders, Drag and Drop) de-duping is enabled.
[LNDEDUPE]	1 or 0	(Required) Controls whether Lotus Notes email de-duping is enabled.
[ESTARTDD]	NUM	(Optional) The beginning day in the email date range search when searching loose PST files.
[ESTARTMM]	NUM	(Optional) The beginning month in the email date range search
[ESTARTYYYY]	NUM	(Optional) The beginning year in the email date range search
[AESTARTDD]	NUM	(Optional) The beginning day in an active email date range search (applies to mounted PST files, Exchange and Public Folders)
[AESTARTMM]	NUM	(Optional) The beginning month in an active email date range search (applies to mounted PST files, Exchange and Public Folders)
[AESTARTYYYY]	NUM	(Optional) The beginning year in an active email date range search (applies to mounted PST files, Exchange and Public Folders)
[EENDDD]	NUM	(Optional) The ending day in the email date range search when searching loose PST files.
[EENDMM]	NUM	(Optional) The ending month in the email date range search when searching loose PST files.
[EENDYYYY]	NUM	(Optional) The ending year in the email date range search when searching loose PST files.

[AEENDDD]	NUM	(Optional) The ending day in an active email date range search (applies to mounted PST files, Exchange and Public Folders)
[AEENDMM]	NUM	(Optional) The ending month in an active email date range search (applies to mounted PST files, Exchange and Public Folders)
[AEENDYYYY]	NUM	(Optional) Then ending year in an active email date range search (applies to mounted PST files, Exchange and Public Folders)
[EATTACHDATES]	1 or 0	(Optional) Indicates whether the *Apply date range search to attachment file dates* box was checked in the loose PST search settings. Not necessary if emails are not being searched or if no date range is defined.
[AEATTACHDATES]	1 or 0	(Optional)) Indicates whether the *Apply date range search to attachment file dates* box was checked in mounted PST, Exchange, and Public Folders search settings.
[EADDYEX]	1 or 0	(Optional) Indicates whether emails with senders or recipients matching the patterns defined for Address/Domain searching in the Loose PST search options will be excluded or included. 0 denotes that the search will hit on only emails to or from the listed addresses or domains. *1* denotes that the search should hit on only emails that do NOT contain the listed addresses or domains. Not necessary if emails are not being searched or if no address/domain filters have been defined.
[AEADDYEX]	1 or 0	(Optional) Indicates whether emails with senders or recipients matching the patterns defined for Address/Domain searching in the mounted PST, Exchange, and Public Folders search options will be excluded or included. 0 denotes that the search will hit only on emails to or from the listed addresses or domains. 1 denotes that the search should hit only on emails that do NOT contain the listed addresses or domains. This option is not necessary if emails are not being searched or if no address/domain filters have been defined.
[LNADDYEX]	1 or 0	(Optional)Indicates whether emails with senders or recipients matching the patterns defined for Address/Domain searching in the Lotus Notes search options will be excluded or included. 0 denotes that the search will hit only on emails to or from the listed addresses or domains. 1 denotes that the search should hit only on emails that do NOT contain the listed addresses or domains. This option is not necessary if emails are not being searched or if no address/domain filters have been defined.
[EPTYPE]	NUM	(Required) Controls which email extraction option is selected in the loose PST search options. 0 = Single target per source. 1 = Collate sources into single target PST. 2 = Generate loose email files from source.
[AEPTYPE]	NUM	(Required) Controls which email extraction option is selected in the Exchange/Mounted PST/Drag and Drop search options. 0 = Single target per source. 1 = Collate sources into single target PST. 2 = Generate loose email files from source.
[EPROCPATH]	text	Optional) Under the loose PST search settings, if you've selected the option to collect email data from multiple PST files ~~or~~ and collate them into a single source, this is the path to the collated PST file.
[AEPROCPATH]	text	(Optional) Under the Exchange/Mounted PST/Drag and Drop search settings, if you've selected the option to collate emails into a single target PST, this is the path to the collated PST.
[EEXPORTFORMAT]	NUM	(Required) Controls which individual email format is selected in the loose PST search settings. 0 = Message files - Unicode (*.msg) 1 = Raw RFC822 (*.eml)
[AEEXPORTFORMAT]	NUM	(Required) Controls which email export format was selected in the Exchange/Mounted PST/Drag and Drop settings. 0 = Message files - Unicode (*.msg) 1 = Raw RFC822 (*.eml)
[LNWORKINGCOPY]	1 or 0	(Required) This option controls whether a working copy of each Lotus Notes NSF file will be created for searching 1 = A working copy will be created prior to copy if able. 0 = The search will be conducted on the original NSF file.
[STOREBEGIN]	text	(Job file only) This entry is written to the job file, followed by the path to the email store, when processing of the email store begins. It is used in the resume feature.
[STOREEND]	text	(Job file only) This entry is written to the job file, followed by the path to the email store, when processing of the email store completes.
[LASTPSTFOLDER]	text	(Job file only) This entry is written to the job file, along with the internal PST path, when a PST folder process begins.
[DETECTENC]	1 or 0	(Required) This value indicates whether or not the *Detect Encrypted Files and Attachments* box was checked in the encryption detection options. 1 = Encryption detection will be performed. 0 = Encryption status will not be determined.

[COPYENCTO]	1 or 0	(Optional) This value indicates whether or not the *Copy encrypted files to a special folder* box was checked in the encryption detection options. Not necessary if [DETECTENC] is 0. 1 = Encrypted files will be copied to the location specified in [ENCTARGET] 0 = Encrypted files will not be copied to a special location
[ENCFULLPATHS]	1 or 0	(Optional) This value indicates whether or not the *Create Full Paths* box was checked in the encryption detection options. Not necessary if either [DETECTENC] or [COPYENCTO] are 0 1 = The full path to the encrypted file will be reflected in the folder structure under the location specified in [ENCTARGET] 0 = Only subfolders will be reflected in the case that [ENCSUBFOLDERS] is 1. No folder structure will be reflected in the case that [ENCSUBFOLDERS] is 0
[ENCROOTFOLDERS]	1 or 0	Optional) This value indicates whether or not the *Create Root Folders* box was checked in the encryption detection options. Not necessary if either [DETECTENC] or [COPYENCTO] are 0 1 = Folders named for the drive letters or UNC servers at the roots of the source paths for encrypted files will be reflected in the path specified in [ENCTARGET] 0 = Drive level and server level folders will not be created in the path specified in [ENCTARGET]
[ENCSUBFOLDERS]	1 or 0	Optional) This value indicates whether or not the *Create Sub Folders* box was checked in the encryption detection options. Not necessary if either [DETECTENC] or [COPYENCTO] are 0 1 = The target will contain subfolders 0 = All files will be written to the same folder with no mirrored structure.
[COPYENCNORMAL]	1 or 0	(Optional) This value indicates whether or not the Copy encrypted files normally box was checked in the encryption detection settings. Not necessary if [DETECTENC] is 0 1 = Files found to be encrypted will be copied to their normal target location 0 = Files found to be encrypted will not be copied to their normal location.
[ENCTARGET]	text	(Optional) This is the alternate path to which encrypted files should be copied. In the OCC file, it may contain variables. In the job file, it will be a fully realized path. Not necessary if either [DETECTENC] or [COPYENCTO] are 0
[NUMSTORES]	NUM	(Job file only) This is the number for PST stores tallied during the run. It is written to the job file at the end of the enumeration phase.
[STORESDONE]	NUM	(Job file only) This is the number of stores completed at the time the job is canceled.
[NUMFILES]	NUM	(Job file only) This is the number of files enumerated. It is recorded at the end of the enumeration phase.
[JOBSIZE]	NUM	(Job file only) This is the size in bytes of the items enumerated for a job. It is recorded after the enumeration phase.
[SOURCEFILECOUNT]	NUM	(Job file only) This is the total number of email source files enumerated. It is written to the job file at the end of the enumeration phase.
[SOURCEBYTECOUNT]	NUM	(Job file only) This is the total size of all enumerated email sources in bytes. It is written to the job file at the end of the enumeration phase.
[ARCHIVEFILECOUNT]	NUM	(Job file only) This is the total number of all archive files enumerated. It is written to the job file at the end of the enumeration phase.
[ARCHIVEBYTECOUNT]	NUM	(Job file only) This is the total size of all enumerated archive files. It is written to the job file at the end of the enumeration phase.
[TCP]	HEX	(Not Manually Editable) This field holds the encrypted password supplied for TrueCrypt volume targets. Manually entered values will not be valid.
[TCEXPATH]	Text	This is the path to the TrueCrypt executable to be used by Harvester. For OCC files, it is automatically determined if you have included the portable version of TrueCrypt in your Harvester directory.
[TCCONTAINER]	Text	This is the path to the TrueCrypt container being mounted during processing.
[TCMOUNTLETTER]	Text	This is the drive letter that the TrueCrypt volume was/will be mounted as during run time. If a letter is not specified or is unavailable, the next available drive letter above G will be used.
[MAILPROCESSOR]	Text	Indicates which mail processor should be used to process a particular email store. Valid values are “ASPOSE”, “REDEMPTION”, and “”
[AOUTMULTIPST]	Text	Indicates whether each email store in an active Outlook-based email collection should be written to its own PST. Valid values are “1” for one PST per mounted source or “0” for a single PST for all active email.

When different job files are created from the same default template, the following common fields could be easily edited and the job saved to a new filename.

[JName] – Custodian name and what appears in job list
[EXTS] – File extensions, definitions or categories

Making modification to just the job name [JName] and resaving to a new filename would keep all other variables (including filtering options) the same. If you need to modify the file types collected, editing the [EXT] field would allow you to make these changes on the fly.

CAUTION:
An improperly formatted job file can prevent a job from running or miss relevant data sources. Be careful and take the time to verify all .occ files.