HARVESTER

HARVESTER PORTABLE CERTIFICATION MANUAL

FEATURES IN HARVESTER

Harvester Portable can be licensed and run from/to a variety of media
(USB flash drives, external hard drives) or a host computer

- ✓ Create job files that can be used to automate collections.
- ✓ Preserve file timestamps and metadata.
- ✓ Maintain the chain of custody.
- ✓ Run Harvester and job files on a variety of media (USB flash drives, external hard drives or internal hard drives).
- ✓ Scriptable and shell out command on job completion.
- ✓ Specify multiple data sources or file lists.
- ✓ Filter collection by file type, date range and file signatures.
- ✓ Copy and preserve long file paths (up to 32,000 characters).
- ✓ Easily resume jobs in the event of a system reboot, crash or other problem.
- ✓ Easily transfer the license to another location.
- ✓ Keyword search loose files, archives, and Microsoft Outlook PSTs and email attachments.
- ✓ Regenerate new filtered PSTs or export to 8 different message formats.
- ✓ De-duplicate loose files and PST messages.
- ✓ “DeNISTing” removes files matching the NSRL (NIST) list.
- ✓ Automatically detect password protected or encrypted files and email attachments.
- ✓ Create full-text searchable indexes and keyword hit reports.
- ✓ Easily access previous job details and reports.
- ✓ VHD utilization allows for creation of file containers
- ✓ VSS allows copying of “in-use” files
- ✓ Process OST’s, and PST’s without using Microsoft Outlook
- ✓ Search and collect email from Microsoft Exchange, Office365, Gmail, Yahoo, Live and other IMAP accounts
- ✓ Microsoft Outlook 32bit and 64bit support
- ✓ 64-bit drag-and-drop support
- ✓ Search and collect files from Google Drive, One Drive, OneDrive for Business, Dropbox, and Box
- ✓ Vera-Crypt utilization allows for creation of encrypted file containers

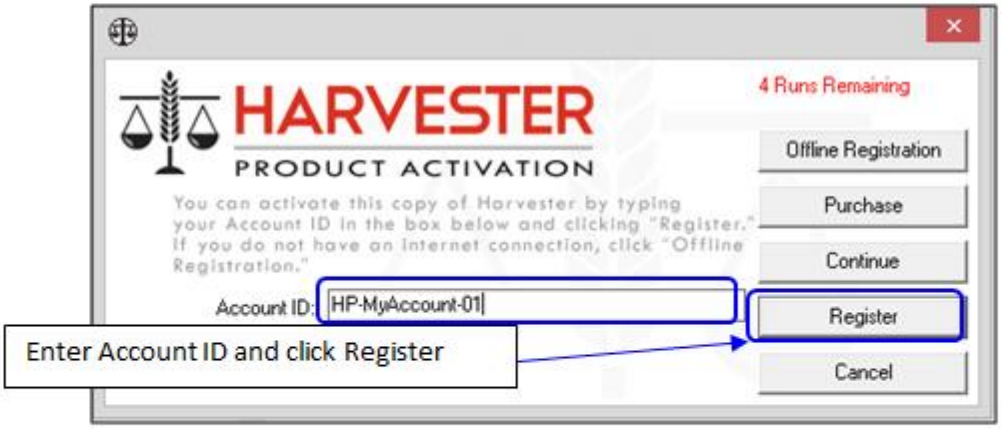
Activate Harvester

The license can easily be deactivated (checked in) and reactivated on a different device when needed.

To download and activate, follow these steps:

Reference Video: **Harvester 5.1 Portable - Download & Activate**

- Download Harvester from the link provided.
- Extract the .zip file contents to your external device (USB flash drive, etc.) or computer.
- Run *Harvester.exe*
- Enter Account ID and click Register to activate the product using online activation



You can activate this copy of Harvester by typing your Account ID in the box below and clicking "Register." If you do not have an internet connection, click "Offline Registration."

Account ID: HP-MyAccount-01

Enter Account ID and click Register

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 is the ability to move it from drive to drive or between different computers.
Once activated on a drive the user can perform a collection, deactivate the license and move the license to another device or computer.
Once you have placed the .zip file contents onto another device or computer, you can activate the license using your same Account ID.

GENERAL TAB

Reference videos:

- Harvester Portable 5.1 Advanced Options -Part 1
- Harvester 5 - Create a Profile OCC

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:
Once saved, the job profiles are stored in the _occ folder by default; however, users can browse to other job locations by clicking ***Open***. With a saved OCC, the Job File indicates the current job profile (.occ) file location.

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.

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.

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

****More about** how to refine your search criteria in much greater detail in the **Data Assessment Mode** section.

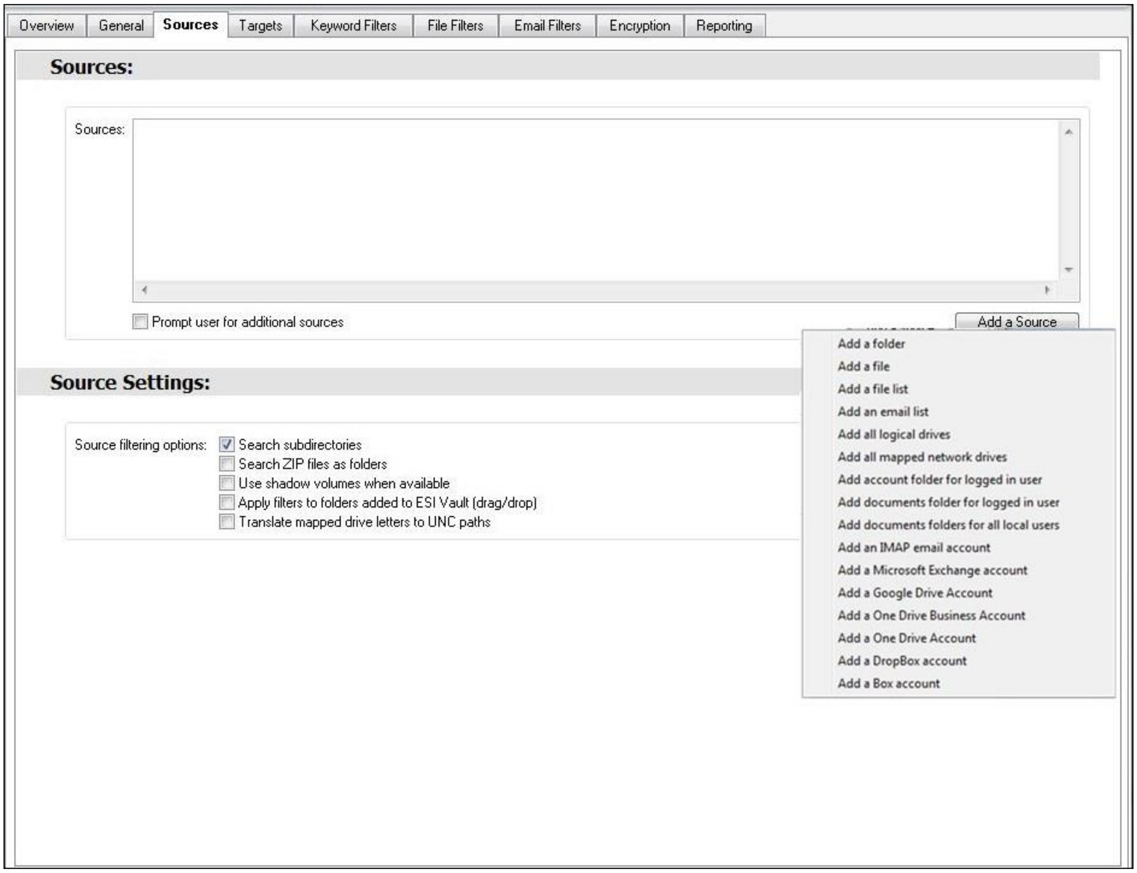
Run in Silent Mode:
Errors that can occur during a project will be logged and this option can often prevent the job from stalling while waiting for a user response (i.e. click the ***Ok*** button).

***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 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:


Overview
General
Sources
Targets
Keyword Filters
File Filters
Email Filters
Encryption
Reporting
Add a Source
Add a folder
Add a file
Add a file list
Add an email list
Add all logical drives
Add all mapped network drives
Add account folder for logged in user
Add documents folder for logged in user
Add documents folders for all local users
Add an IMAP email account
Add a Microsoft Exchange account
Add a Google Drive Account
Add a One Drive Business Account
Add a One Drive Account
Add a DropBox account
Add a Box account

Sources:

Sources:

☐ Prompt user for additional sources

Source Settings:

Source filtering options:
☑ Search subdirectories
☐ Search ZIP files as folders
☐ Use shadow volumes when available
☐ Apply filters to folders added to ESI Vault (drag/drop)
☐ Translate mapped drive letters to UNC paths

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 target because the operating system does not check to see if the target exists. Harvester does not treat symbolic links as folders or files due to the possibility that

- 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 only field in the text file. The list file can contain one file path or directory per line.

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:

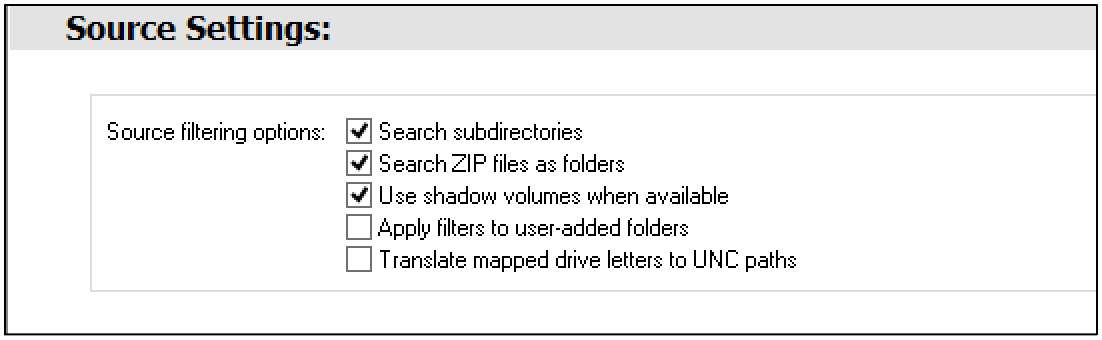
- **[LDrive]** This variable (with the brackets) indicates that the program should search all logical drives that are connected to the computer. This includes flash drives, CDs, internal hard drives and RAID devices. It does not include network shares or the device that the Harvester software is running 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: (\netsharefilesjohndoe1)). 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.

SOURCE SETTINGS:



Source Settings:

Source filtering options:
☑ Search subdirectories
☑ Search ZIP files as folders
☑ Use shadow volumes when available
□ Apply filters to user-added folders
□ Translate mapped drive letters to UNC paths

Search Subdirectories: This option specifies whether subdirectories are searched. Deselecting this option will cause to 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 and extension, date, 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.

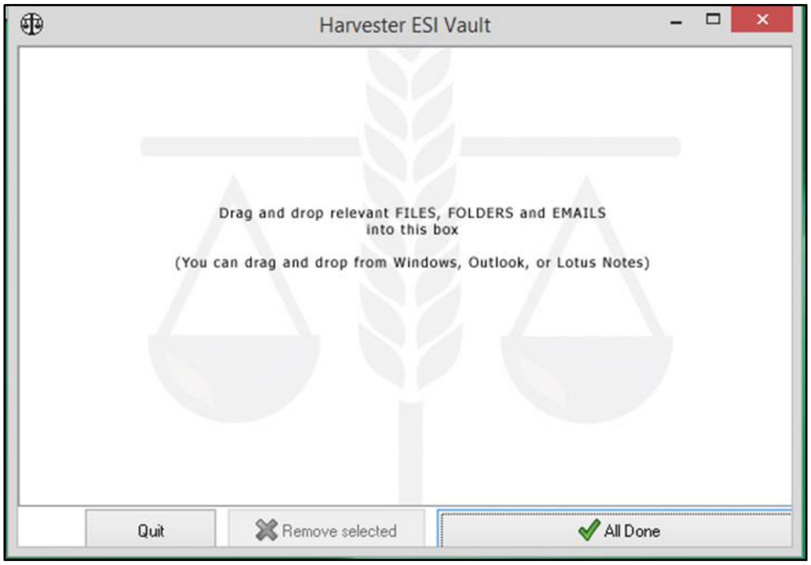
Shadow Copy – also known as Volume Shadow Copy Service or VSS – is a technology included in Microsoft Windows that allows taking manual or automatic backup copies or snapshots of computer files or volumes, even when they are in use. It is implemented as a Windows service called the Volume Shadow Copy 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, removable, or network volumes by any Windows component that uses this technology, such as when creating a scheduled Windows Backup or automatic System Restore point.

Note: Most remote network volumes do not support shadow copies due to a lack of low-level access by the executing machine.

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 mapped network drives paths to their UNC paths. The UNC path and file name will appear in the *filelist.txt* and *folderlist.txt* in the log folder.

ESI "Easy" Vault is commonly used when distributing self-collection kits or jobs launched from Harvester Server for legal hold notice. By providing this interface and instructions specific to each job, custodians can easily identify items relevant to a matter without violating international privacy laws. The ESI Vault interface is a window that supports dragging and dropping of the following types of items.



Drag and drop relevant FILES, FOLDERS and EMAILS
into this box

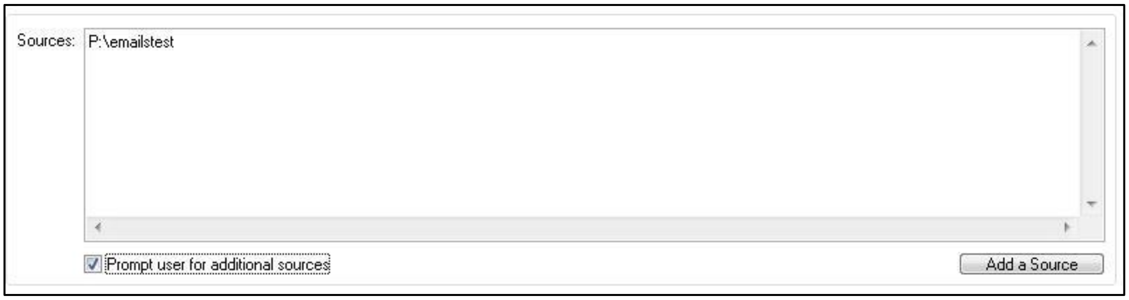
(You can drag and drop from Windows, Outlook, or Lotus Notes)

- **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.

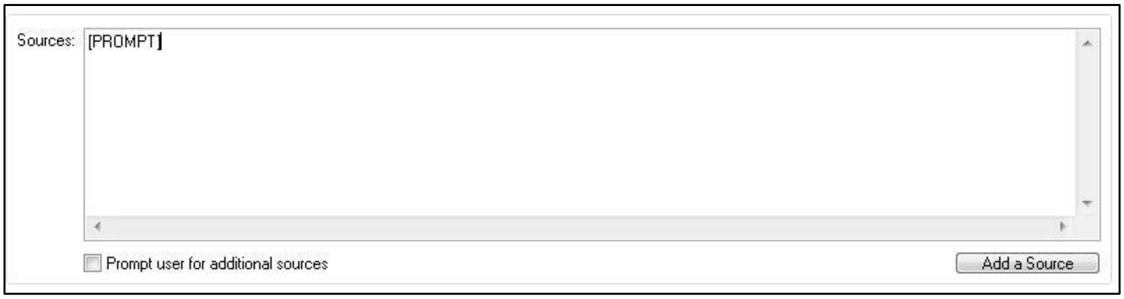
There are two scenarios that will launch the ESI Vault during a job:

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

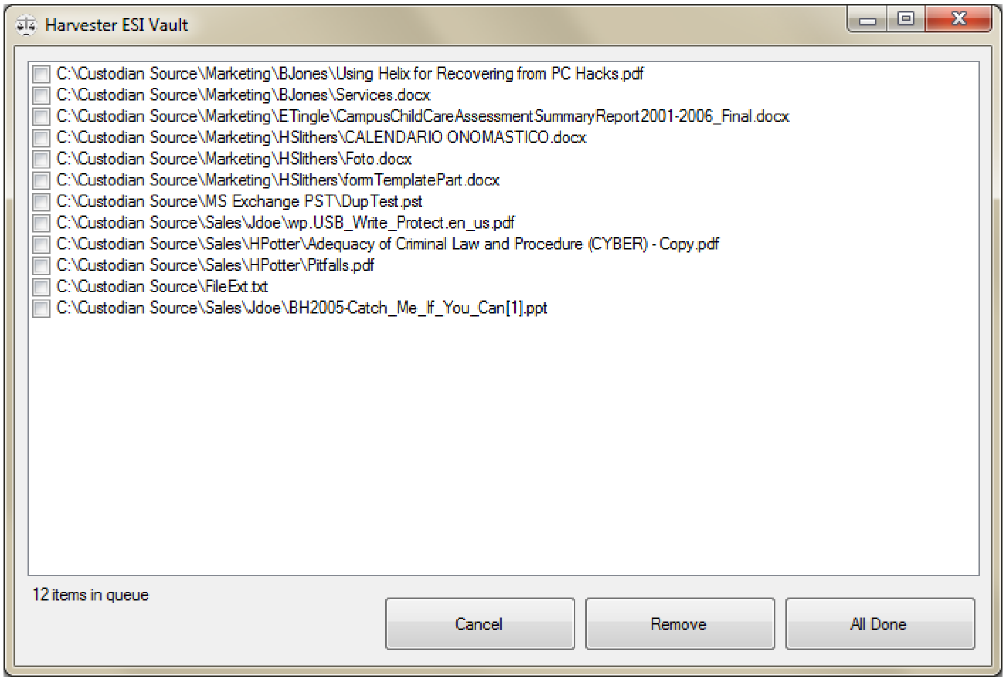

Sources: P:\emailstest

☑ Prompt user for additional sources

- ***One or more sources is set as [PROMPT]***


Sources: [PROMPT]

☐ Prompt user for additional sources



□ C:\Custodian Source\Marketing\BJones\Using Helix for Recovering from PC Hacks.pdf
□ C:\Custodian Source\Marketing\BJones\Services.docx
□ C:\Custodian Source\Marketing\ETingle\CampusChildCareAssessmentSummaryReport2001-2006_Final.docx
□ C:\Custodian Source\Marketing\HSlithers\CALENDARIO ONOMASTICO.docx
□ C:\Custodian Source\Marketing\HSlithers\Foto.docx
□ C:\Custodian Source\Marketing\HSlithers\formTemplatePart.docx
□ C:\Custodian Source\MS Exchange PST\DupTest.pst
□ C:\Custodian Source\Sales\Jdoe\wp.USB_Write_Protect.en_us.pdf
□ C:\Custodian Source\Sales\HPotter\Adequacy of Criminal Law and Procedure (CYBER) - Copy.pdf
□ C:\Custodian Source\Sales\HPotter\Pitfalls.pdf
□ C:\Custodian Source\FileExt.txt
□ C:\Custodian Source\Sales\Jdoe\BH2005-Catch_Me_If_You_Can[1].ppt

12 items in queue

- When adding items to the ESI Vault window, the window displays a list and the number of items added.
- Any sources that the user adds can be removed by selecting the items and clicking the ***Remove*** 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.

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 below.

EXCH 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

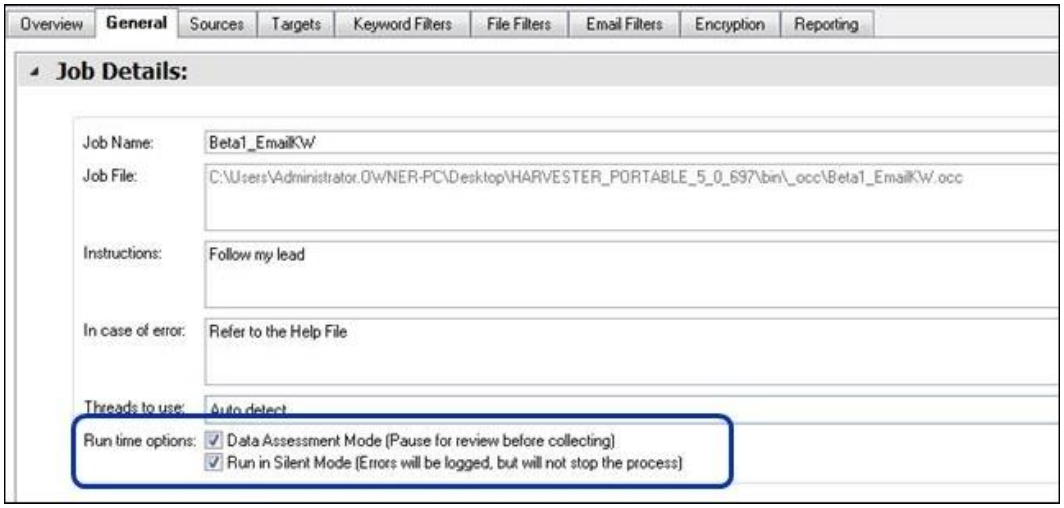
Cloud source connections allow users to enumerate and collect loose files directly from web based cloud storage accounts such as Google Drive, OneDrive, OneDrive for Business, Box, and Dropbox.

Data Assessment Mode

Harvester runs in two separate stages:

1. **Enumeration** (inventory stage) This is the first phase where it goes through the specified sources and records which files meet filter criteria.
2. **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.



Job Details:

Job Name: Beta1_EmailKW

Job File: C:\Users\Administrator.OWNER-PC\Desktop\HARVESTER_PORTABLE_5_0_697\bin_occ\Beta1_EmailKW.occ

Instructions: Follow my lead

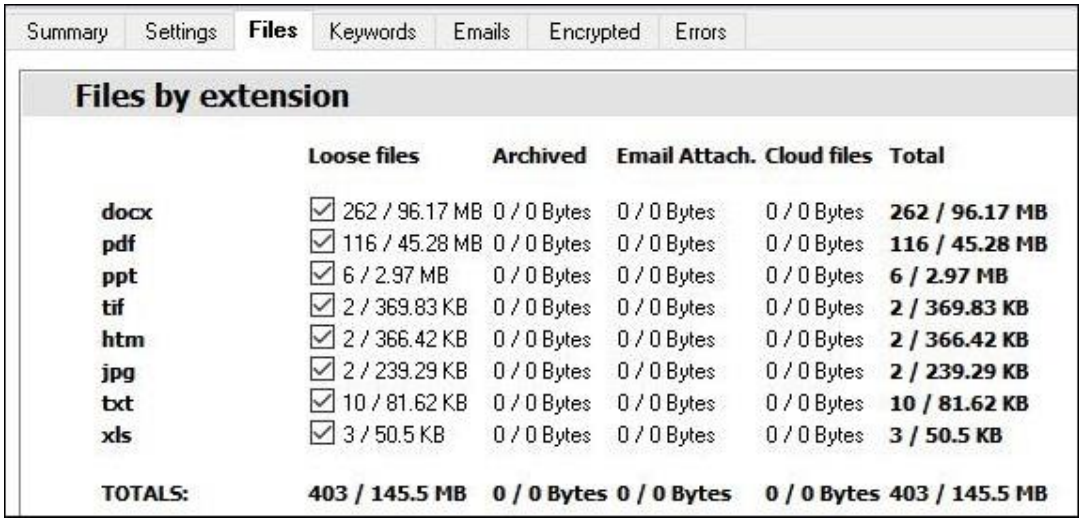
In case of error: Refer to the Help File

Threads to use: Auto detect

Run time options: ☑ Data Assessment Mode (Pause for review before collecting)
☑ Run in Silent Mode (Errors will be logged, but will not stop the process)

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.

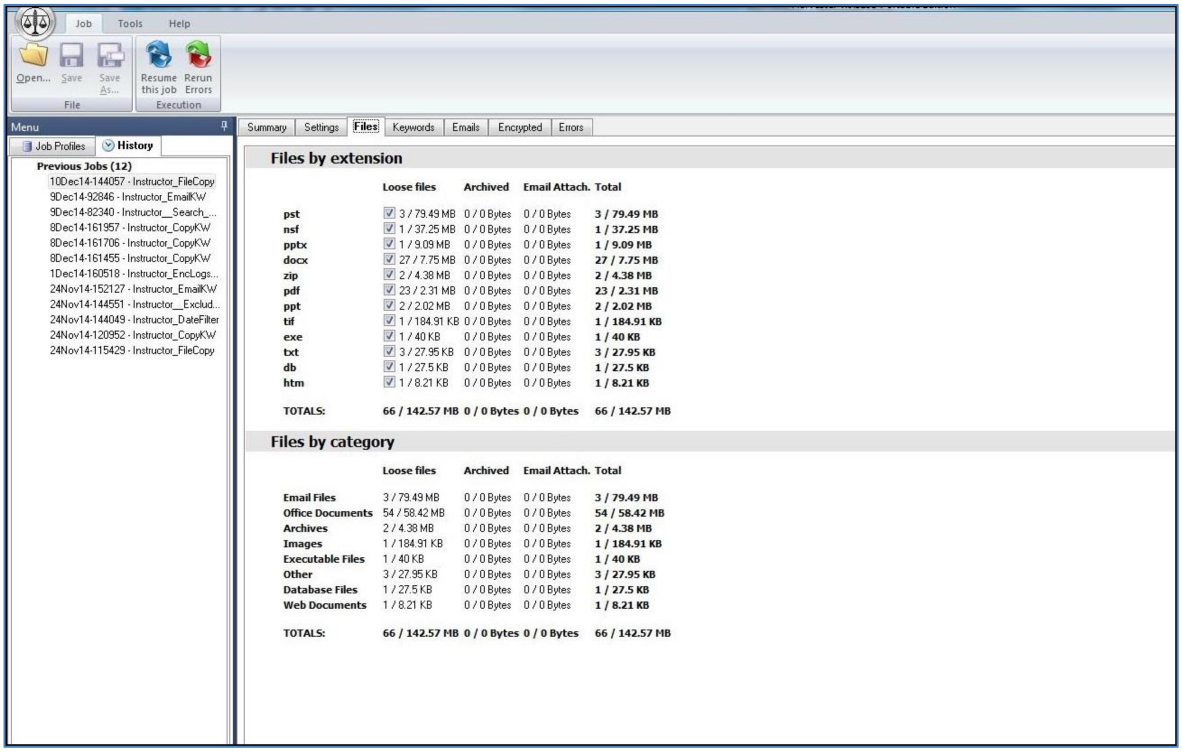
Under the Files tab, check for unnecessary file extensions. Choose and uncheck unnecessary extensions before resuming into the copy phase.



Files by extension

	Loose files	Archived	Email Attach.	Cloud files	Total
docx	☑ 262 / 96.17 MB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**262 / 96.17 MB**
pdf	☑ 116 / 45.28 MB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**116 / 45.28 MB**
ppt	☑ 6 / 2.97 MB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**6 / 2.97 MB**
tif	☑ 2 / 369.83 KB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**2 / 369.83 KB**
htm	☑ 2 / 366.42 KB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**2 / 366.42 KB**
jpg	☑ 2 / 239.29 KB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**2 / 239.29 KB**
txt	☑ 10 / 81.62 KB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**10 / 81.62 KB**
xls	☑ 3 / 50.5 KB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	**3 / 50.5 KB**
TOTALS:	**403 / 145.5 MB**	**0 / 0 Bytes**	**0 / 0 Bytes**	**0 / 0 Bytes**	**403 / 145.5 MB**

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



Previous Jobs (12)

10Dec14-144057 - Instructor_FileCopy
9Dec14-92846 - Instructor_EmailKW
9Dec14-82340 - Instructor__Search_...
8Dec14-161957 - Instructor_CopyKW
8Dec14-161706 - Instructor_CopyKW
8Dec14-161455 - Instructor_CopyKW
1Dec14-160518 - Instructor_EncLogs...
24Nov14-152127 - Instructor_EmailKW
24Nov14-144551 - Instructor__Exclud...
24Nov14-144049 - Instructor_DateFilter
24Nov14-120952 - Instructor_CopyKW
24Nov14-115429 - Instructor_FileCopy

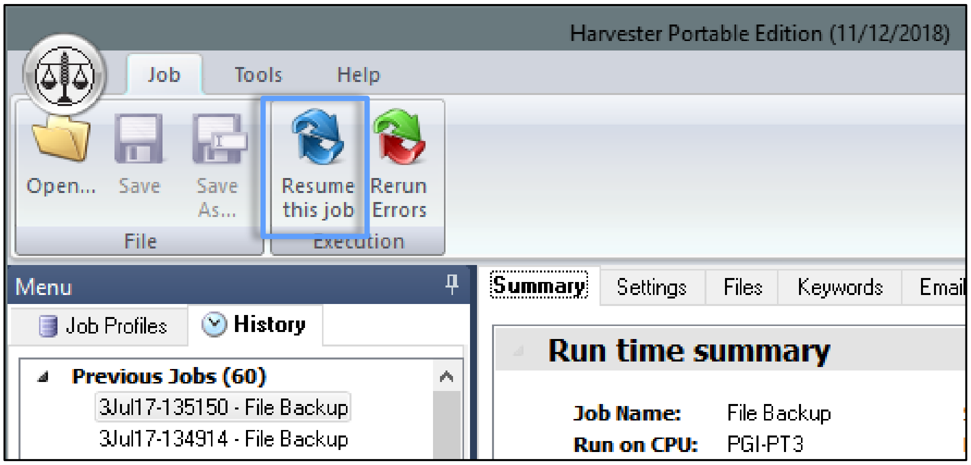
Files by extension

	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**

Files by category

	Loose files	Archived	Email Attach.	Total
Email Files	3 / 79.49 MB	0 / 0 Bytes	0 / 0 Bytes	**3 / 79.49 MB**
Office Documents	54 / 58.42 MB	0 / 0 Bytes	0 / 0 Bytes	**54 / 58.42 MB**
Archives	2 / 4.38 MB	0 / 0 Bytes	0 / 0 Bytes	**2 / 4.38 MB**
Images	1 / 184.91 KB	0 / 0 Bytes	0 / 0 Bytes	**1 / 184.91 KB**
Executable Files	1 / 40 KB	0 / 0 Bytes	0 / 0 Bytes	**1 / 40 KB**
Other	3 / 27.95 KB	0 / 0 Bytes	0 / 0 Bytes	**3 / 27.95 KB**
Database Files	1 / 27.5 KB	0 / 0 Bytes	0 / 0 Bytes	**1 / 27.5 KB**
Web Documents	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**

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



Job Name: File Backup
Run on CPU: PGI-PT3

To Resume a job later (Collect the identified data):

If the logs are available but the job is not listed under the History tab, select Job Profiles > Resume an Incomplete Job. Browse to the ***_jobfile.scj*** for the job you wish to resume and click ***Open***. The *_jobfile.scj* will be in the logs folder. You can also select ***Open*** from the **File** ribbon and navigate to the logs folder. After selecting the logs folder, the History tab will update with the selected job and the ***Resume this job*** icon will be available

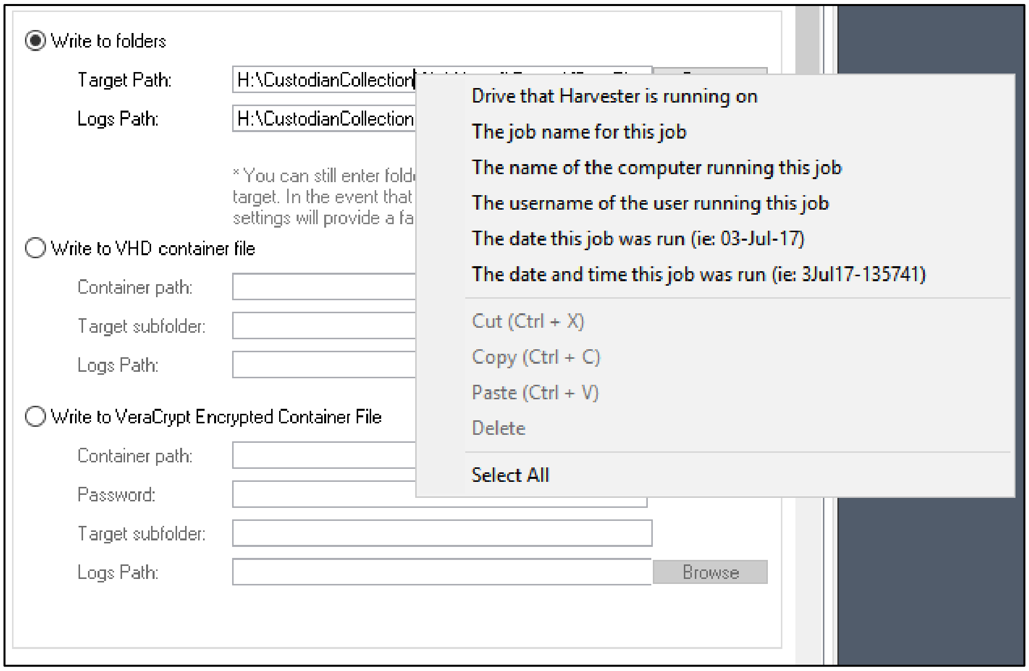
Targets Tab

TARGET PATH

Harvester Portable 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.



● Write to folders

Target Path: H:\CustodianCollection

Logs Path: H:\CustodianCollection

* You can still enter fold
target. In the event that
settings will provide a fa

○ Write to VHD container file

Container path:

Target subfolder:

Logs Path:

○ Write to VeraCrypt Encrypted Container File

Container path:

Password:

Target subfolder:

Logs Path:

Drive that Harvester is running on
The job name for this job
The name of the computer running this job
The username of the user running this job
The date this job was run (ie: 03-Jul-17)
The date and time this job was run (ie: 3Jul17-135741)
Cut (Ctrl + X)
Copy (Ctrl + C)
Paste (Ctrl + V)
Delete
Select All
Browse

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.

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.

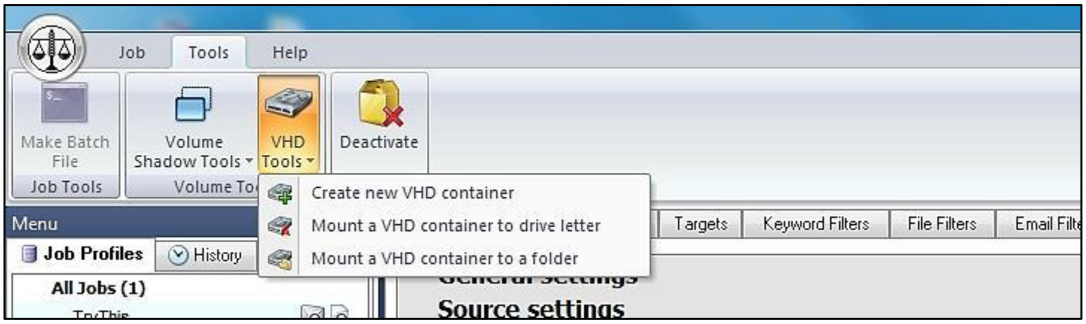
VHD creation is *automatic* when ***Write to VHD file container*** is used as a target.

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.

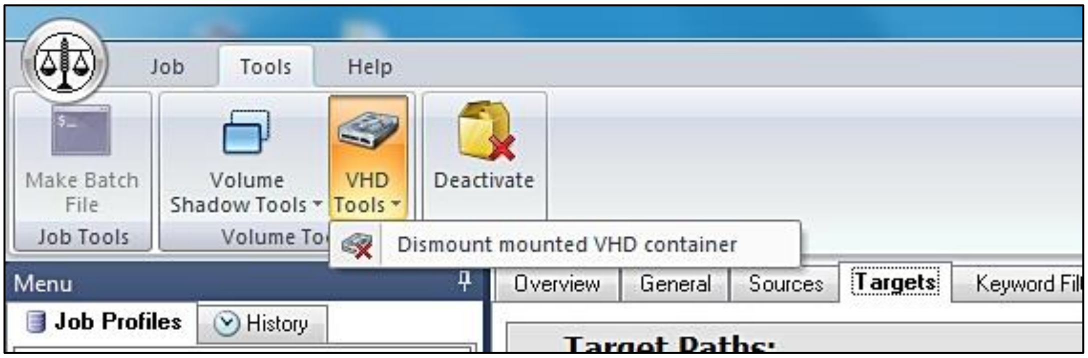
When choosing to copy files to a VHD container, Harvester creates a VHD file at the location specified and formats it like a hard drive. As the Harvester job enters the copy phase, the files are written to the VHD container instead of a target folder. After the job has run, the VHD file itself can later 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*.



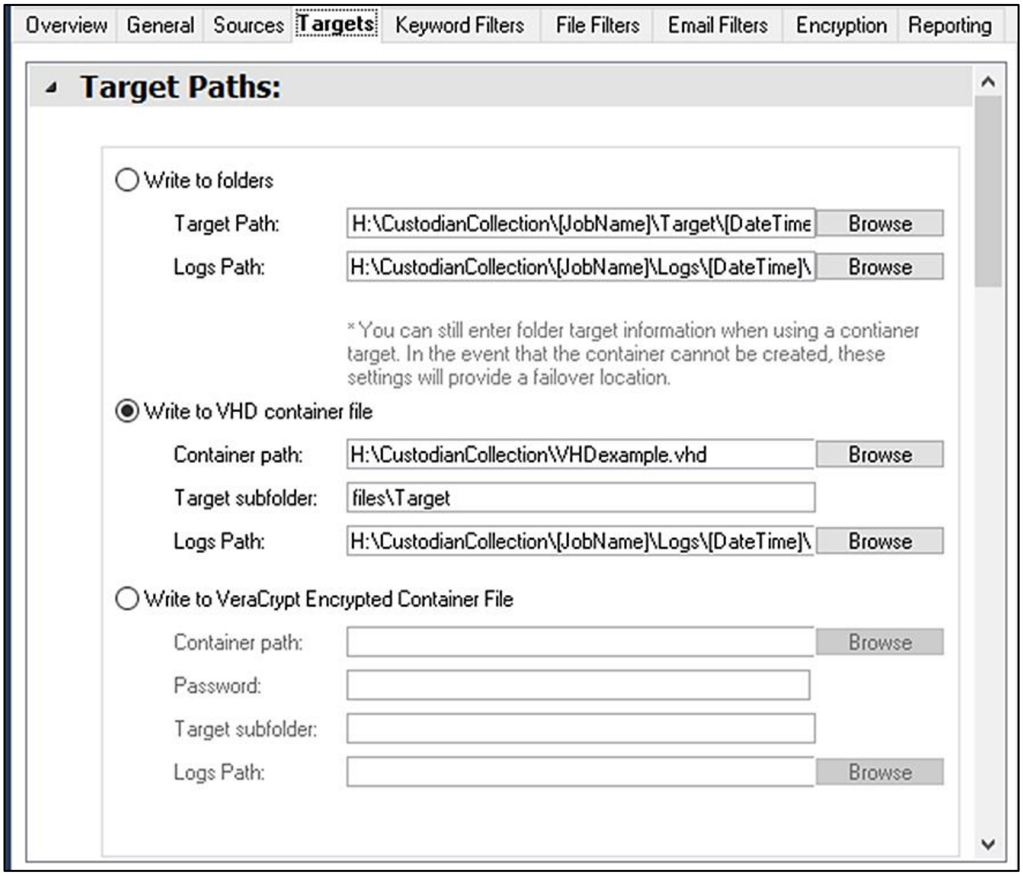
After a VHD is mounted, the option to dismount a VHD container is available. VHD containers are automatically dismounted and detached at the end of the job.



Under Target path,

- ***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 to any preferred location.

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



Target Paths:

○ Write to folders

Target Path: H:\CustodianCollection\[JobName]\Target\[DateTime

Logs Path: H:\CustodianCollection\[JobName]\Logs\[DateTime]\

*You can still enter folder target information when using a contianer target. In the event that the container cannot be created, these settings will provide a failover location.

● Write to VHD container file

Container path: H:\CustodianCollection\VHDexample.vhd

Target subfolder: files\Target

Logs Path: H:\CustodianCollection\[JobName]\Logs\[DateTime]\

○ Write to VeraCrypt Encrypted Container File

Container path:

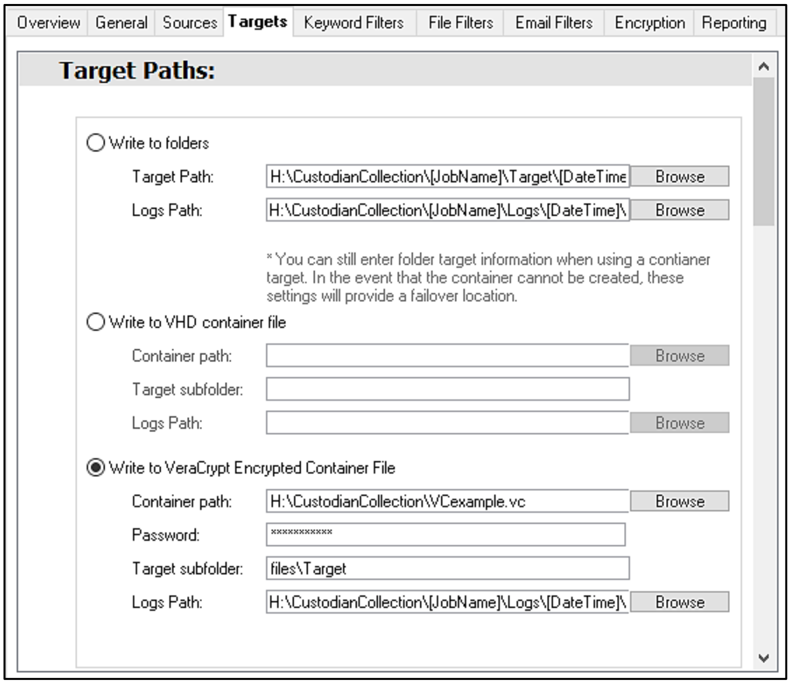
Password:

Target subfolder:

Logs Path:

WRITE TO VERACRYPT ENCRYPTED 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.



Target Paths:

○ Write to folders

Target Path: H:\CustodianCollection\[JobName]\Target\[DateTime

Logs Path: H:\CustodianCollection\[JobName]\Logs\[DateTime]\

* You can still enter folder target information when using a contianer target. In the event that the container cannot be created, these settings will provide a failover location.

○ Write to VHD container file

Container path:

Target subfolder:

Logs Path:

● Write to VeraCrypt Encrypted Container File

Container path: H:\CustodianCollection\VCexample.vc

Password: ************

Target subfolder: files\Target

Logs Path: H:\CustodianCollection\[JobName]\Logs\[DateTime]\

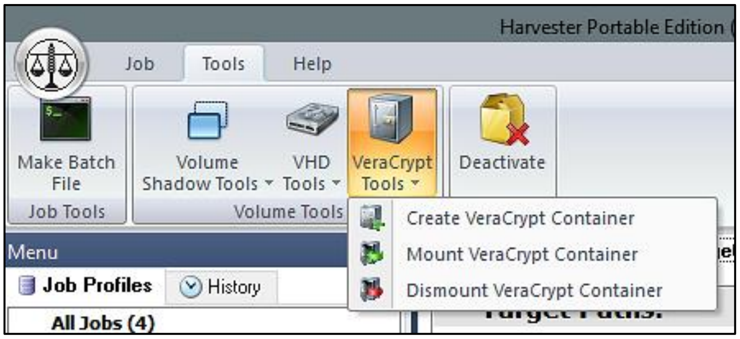
VeraCrypt container creation is *automatic* when *Write to VeraCrypt encrypted container file* is used as a target.

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 mounted to a drive letter 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 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.

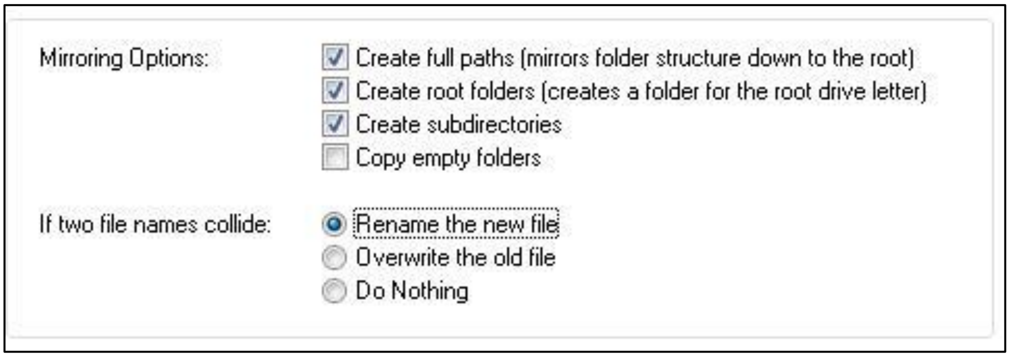
MIRRORING OPTIONS

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, 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.



Mirroring Options:
☑ Create full paths (mirrors folder structure down to the root)
☑ Create root folders (creates a folder for the root drive letter)
☑ Create subdirectories
□ Copy empty folders

If two file names collide:
● Rename the new file
○ Overwrite the old file
○ Do Nothing

If files collide:

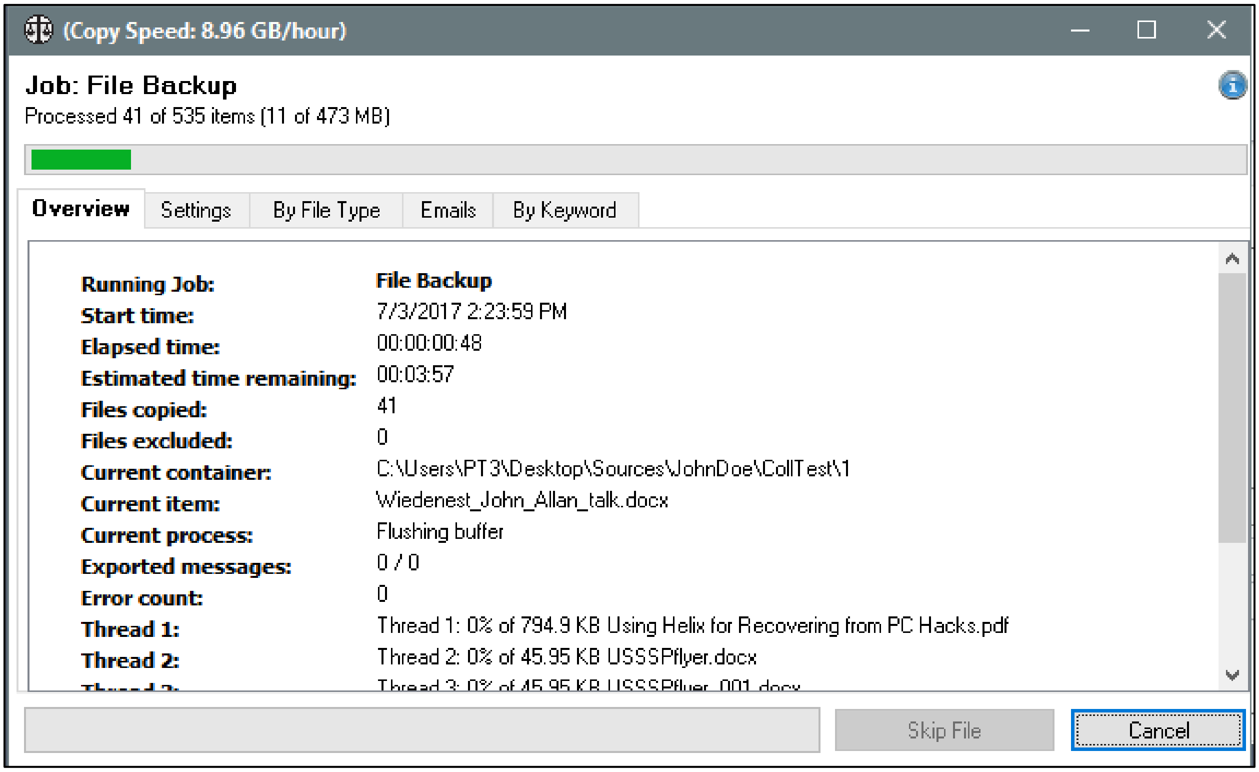
- **Do Nothing:**
  This ignores any files that already exist at the destination and does not include their counterparts in the source directory as responsive.
- **Overwrite Existing Files:**
  This option forces any files that already exist in the target folder to be overwritten.
- **Rename Files on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

Job Progress Console

RUNNING A HARVESTER JOB

To run a Harvester job click on the profile in the list and click ***Run***.

A confirmation window will appear that requires the user to click ***Ok*** to continue. Once the job begins, the Harvester progress console will appear and provide important feedback as well as valuable real-time statistics**.**



(Copy Speed: 8.96 GB/hour)

Job: File Backup

Processed 41 of 535 items (11 of 473 MB)

Running Job: File Backup
Start time: 7/3/2017 2:23:59 PM
Elapsed time: 00:00:00:48
Estimated time remaining: 00:03:57
Files copied: 41
Files excluded: 0
Current container: C:\Users\PT3\Desktop\Sources\JohnDoe\CollTest\1
Current item: Wiedenest_John_Allan_talk.docx
Current process: Flushing buffer
Exported messages: 0 / 0
Error count: 0
Thread 1: Thread 1: 0% of 794.9 KB Using Helix for Recovering from PC Hacks.pdf
Thread 2: Thread 2: 0% of 45.95 KB USSSPflyer.docx
Thread 3: Thread 3: 0% of 45.95 KB USSSPflyer_001.docx

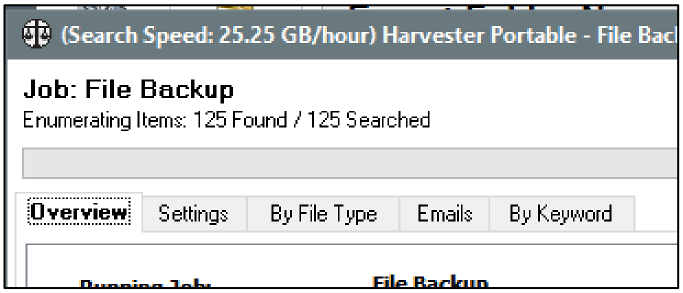
The first tab in the job console displays several useful statistics including:

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

HARVESTER RUNS IN TWO PHASES:

- **Enumeration**

In the progress window Harvester displays ***Enumerating Items***. This is the first phase where it goes through the specified sources and records which files and emails meet filter criteria. In the example below, during enumeration Harvester has found 200 responsive files out of 215 files searched (so far).



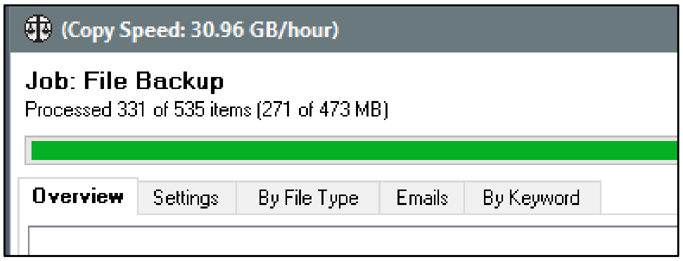
(Search Speed: 25.25 GB/hour) Harvester Portable - File Bac

Job: File Backup

Enumerating Items: 125 Found / 125 Searched

- **Processing**

When enumeration has completed, Harvester will display ***Processed*** which is the copy phase. It then goes through the list it made during enumeration, and copies the items and hash verifies the copies.

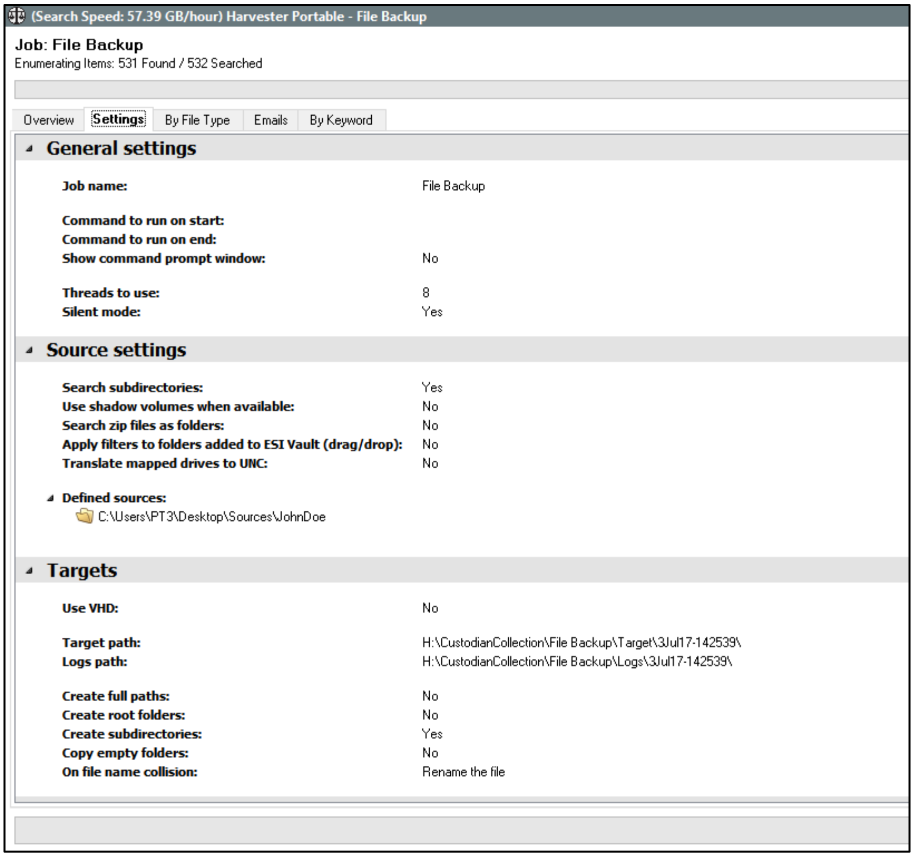


(Copy Speed: 30.96 GB/hour)

Job: File Backup

Processed 331 of 535 items (271 of 473 MB)

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.



(Search Speed: 57.39 GB/hour) Harvester Portable - File Backup

Job: File Backup

Enumerating Items: 531 Found / 532 Searched

General settings

Job name:	File Backup
Command to run on start:	
Command to run on end:	
Show command prompt window:	No
Threads to use:	8
Silent mode:	Yes

Source settings

Search subdirectories:	Yes
Use shadow volumes when available:	No
Search zip files as folders:	No
Apply filters to folders added to ESI Vault (drag/drop):	No
Translate mapped drives to UNC:	No

Defined sources:

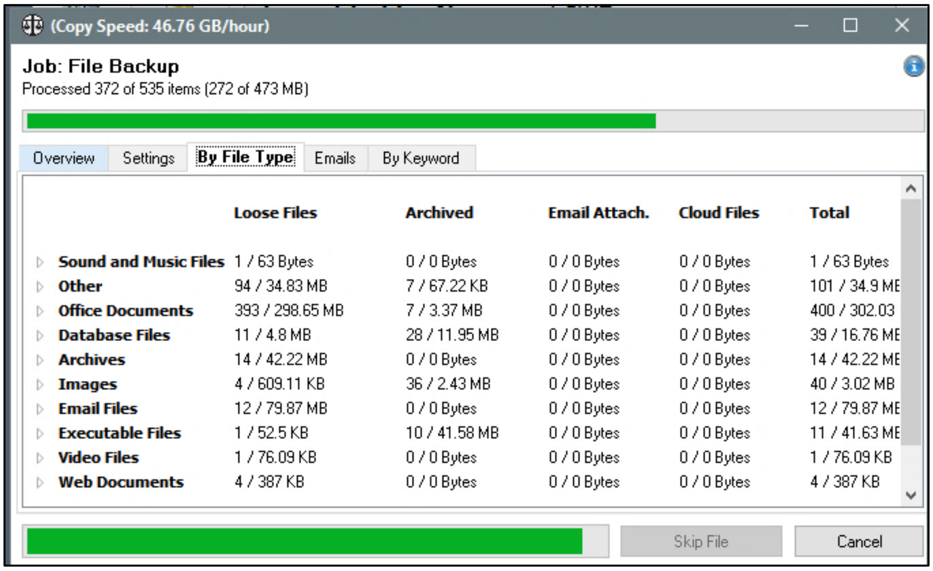
C:\Users\PT3\Desktop\Sources\JohnDoe

Targets

Use VHD:	No
Target path:	H:\CustodianCollection\File Backup\Target\3Jul17-142539\
Logs path:	H:\CustodianCollection\File Backup\Logs\3Jul17-142539\
Create full paths:	No
Create root folders:	No
Create subdirectories:	Yes
Copy empty folders:	No
On file name collision:	Rename the file

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***.



(Copy Speed: 46.76 GB/hour)

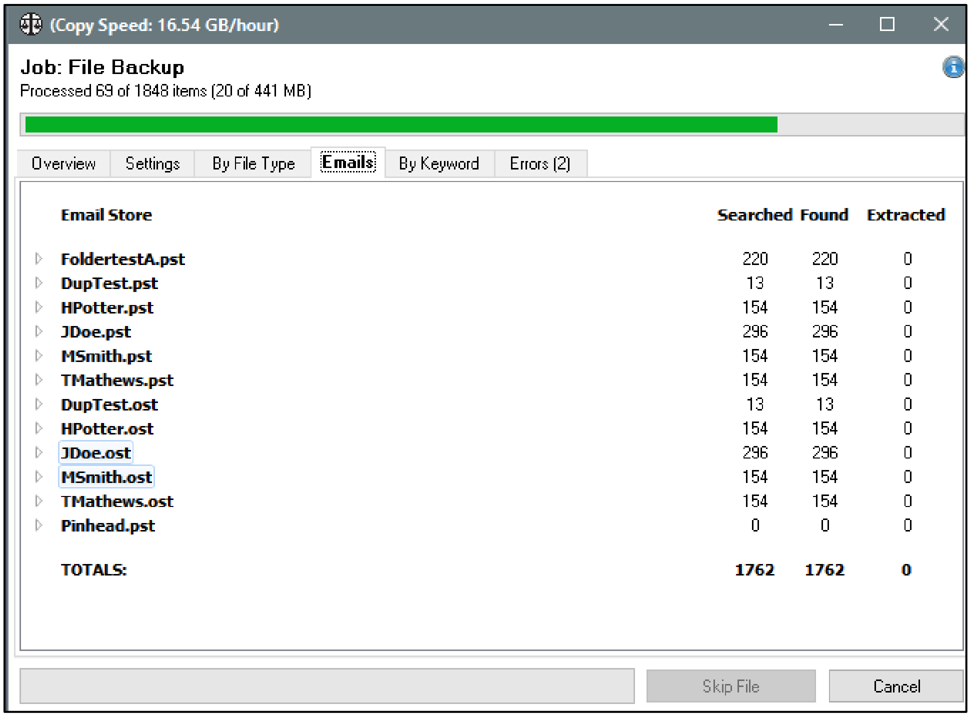
Job: File Backup

Processed 372 of 535 items (272 of 473 MB)

	Loose Files	Archived	Email Attach.	Cloud Files	Total
Sound and Music Files	1 / 63 Bytes	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	1 / 63 Bytes
Other	94 / 34.83 MB	7 / 67.22 KB	0 / 0 Bytes	0 / 0 Bytes	101 / 34.9 ME
Office Documents	393 / 298.65 MB	7 / 3.37 MB	0 / 0 Bytes	0 / 0 Bytes	400 / 302.03
Database Files	11 / 4.8 MB	28 / 11.95 MB	0 / 0 Bytes	0 / 0 Bytes	39 / 16.76 ME
Archives	14 / 42.22 MB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	14 / 42.22 ME
Images	4 / 609.11 KB	36 / 2.43 MB	0 / 0 Bytes	0 / 0 Bytes	40 / 3.02 MB
Email Files	12 / 79.87 MB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	12 / 79.87 ME
Executable Files	1 / 52.5 KB	10 / 41.58 MB	0 / 0 Bytes	0 / 0 Bytes	11 / 41.63 ME
Video Files	1 / 76.09 KB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	1 / 76.09 KB
Web Documents	4 / 387 KB	0 / 0 Bytes	0 / 0 Bytes	0 / 0 Bytes	4 / 387 KB

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.

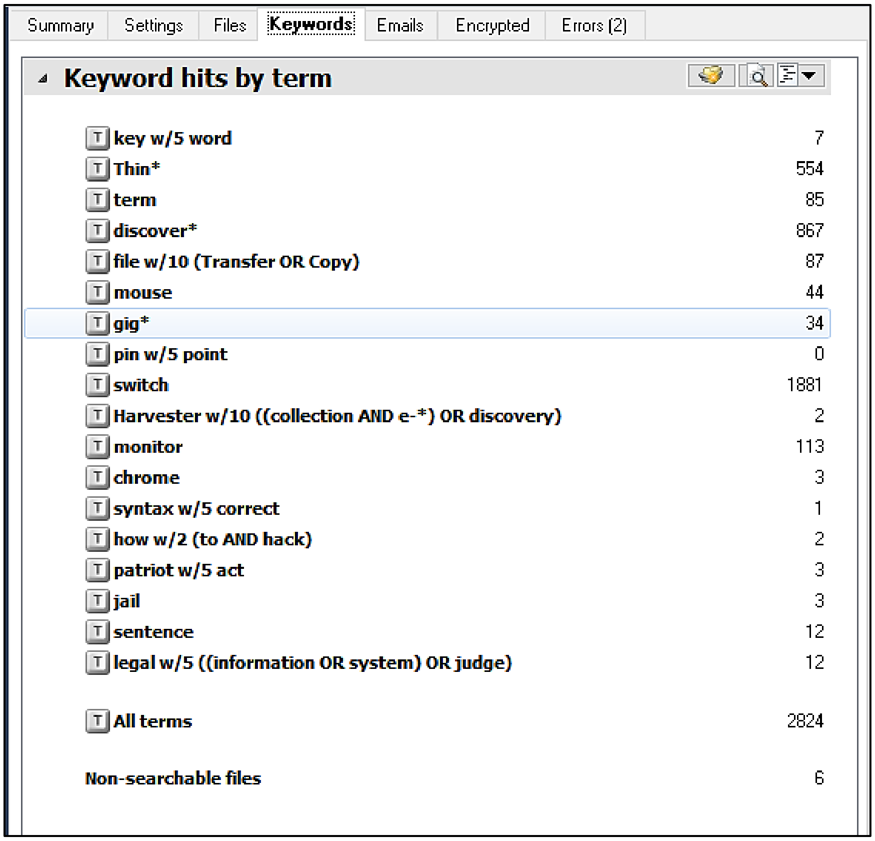


Job: File Backup
Processed 69 of 1848 items (20 of 441 MB)

Email Store	Searched	Found	Extracted
FoldertestA.pst	220	220	0
DupTest.pst	13	13	0
HPotter.pst	154	154	0
JDoe.pst	296	296	0
MSmith.pst	154	154	0
TMathews.pst	154	154	0
DupTest.ost	13	13	0
HPotter.ost	154	154	0
JDoe.ost	296	296	0
MSmith.ost	154	154	0
TMathews.ost	154	154	0
Pinhead.pst	0	0	0
TOTALS:	1762	1762	0

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.**



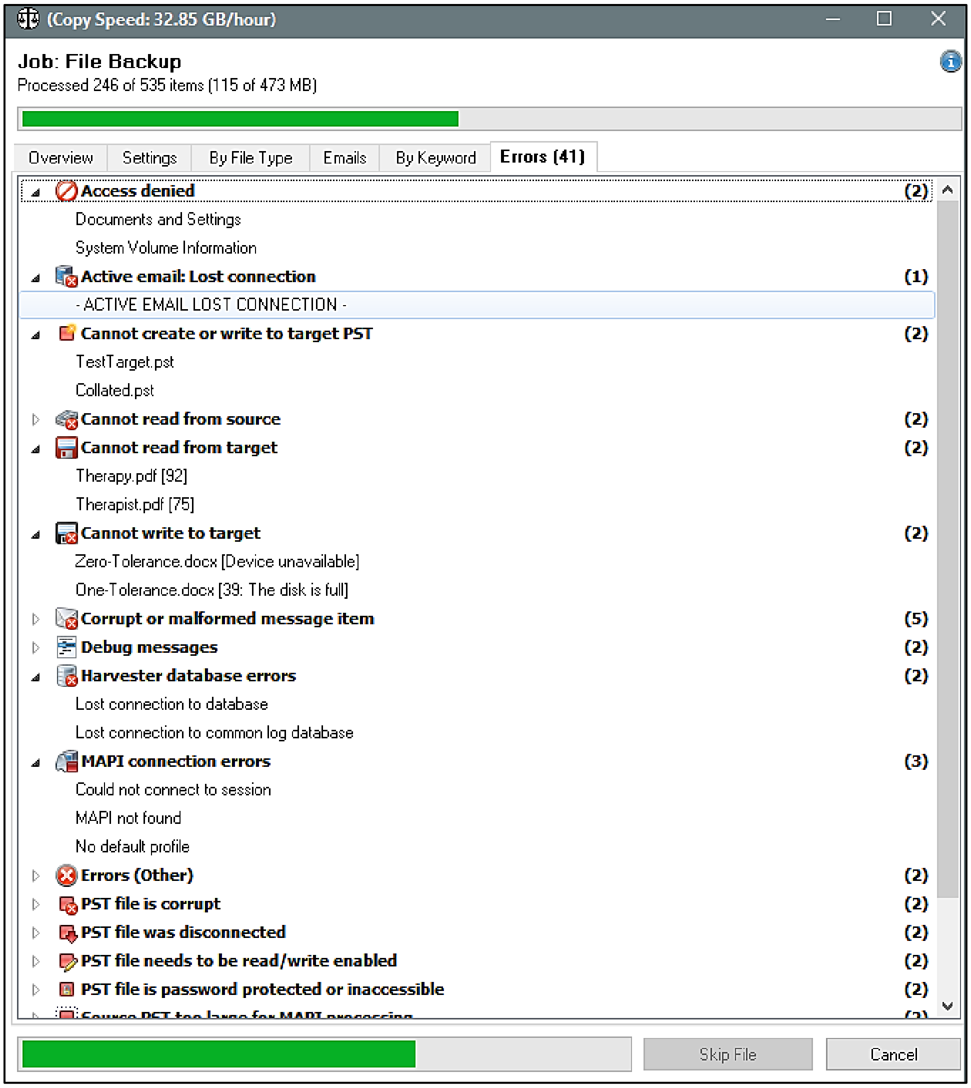
Keyword hits by term

key w/5 word	7
Thin*	554
term	85
discover*	867
file w/10 (Transfer OR Copy)	87
mouse	44
gig*	34
pin w/5 point	0
switch	1881
Harvester w/10 ((collection AND e-*) OR discovery)	2
monitor	113
chrome	3
syntax w/5 correct	1
how w/2 (to AND hack)	2
patriot w/5 act	3
jail	3
sentence	12
legal w/5 ((information OR system) OR judge)	12
All terms	2824
Non-searchable files	6

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 if it is installed.



(Copy Speed: 32.85 GB/hour)

Job: File Backup

Processed 246 of 535 items (115 of 473 MB)

```
Access denied                                        (2)
    Documents and Settings
    System Volume Information
Active email: Lost connection                        (1)
    - ACTIVE EMAIL LOST CONNECTION -
Cannot create or write to target PST                 (2)
    TestTarget.pst
    Collated.pst
Cannot read from source                              (2)
Cannot read from target                              (2)
    Therapy.pdf [92]
    Therapist.pdf [75]
Cannot write to target                               (2)
    Zero-Tolerance.docx [Device unavailable]
    One-Tolerance.docx [39: The disk is full]
Corrupt or malformed message item                    (5)
Debug messages                                       (2)
Harvester database errors                            (2)
    Lost connection to database
    Lost connection to common log database
MAPI connection errors                               (3)
    Could not connect to session
    MAPI not found
    No default profile
Errors (Other)                                       (2)
PST file is corrupt                                  (2)
PST file was disconnected                            (2)
PST file needs to be read/write enabled              (2)
PST file is password protected or inaccessible       (2)
Source PST too large for MAPI processing             (2)
```

Keyword Filters

Reference video: Harvester Portable 5.1 Advanced Options Part 3

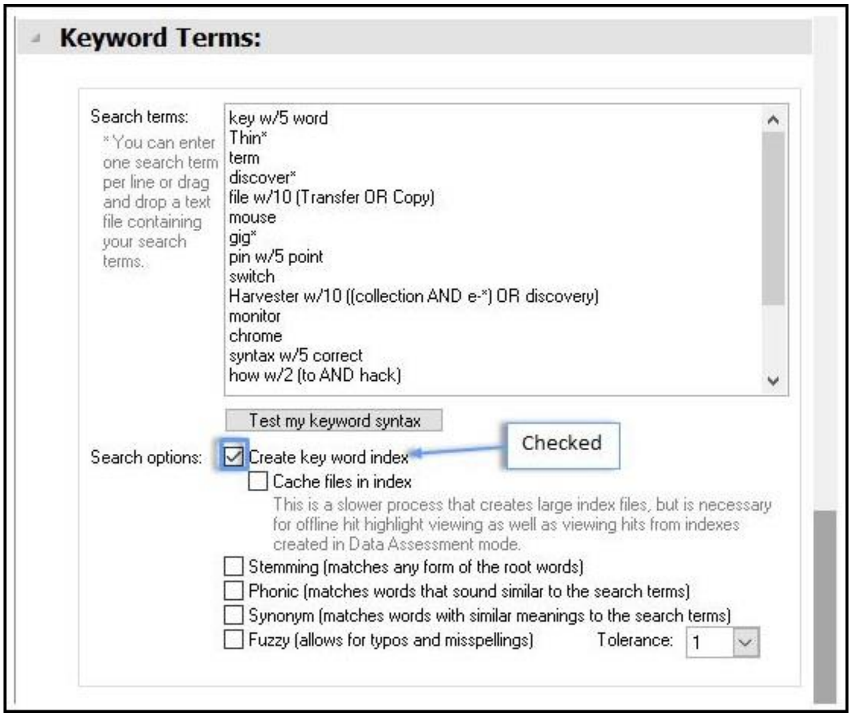
Keyword filtering is one of the most commonly used Harvester features. It is in this group of settings that users can perform targeted e-Discovery collections and filtering processes.

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.

CACHE INDEXING

When *Create key word Index* and *Data Assessment* options are selected



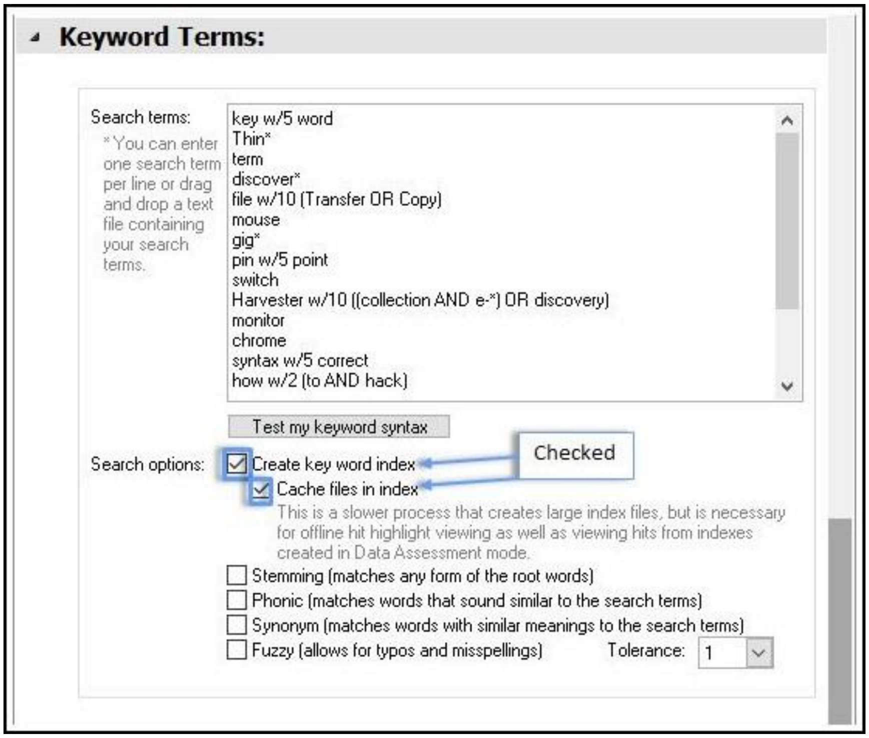
Keyword Terms:

Search terms: *You can enter one search term per line or drag and drop a text file containing your search terms.

```
key w/5 word
Thin*
term
discover*
file w/10 (Transfer OR Copy)
mouse
gig*
pin w/5 point
switch
Harvester w/10 ((collection AND e-*) OR discovery)
monitor
chrome
syntax w/5 correct
how w/2 (to AND hack)
```

Search options:
- ☑ Create key word index
  - ☐ Cache files in index
    This is a slower process that creates large index files, but is necessary for offline hit highlight viewing as well as viewing hits from indexes created in Data Assessment mode.
- ☐ Stemming (matches any form of the root words)
- ☐ Phonic (matches words that sound similar to the search terms)
- ☐ Synonym (matches words with similar meanings to the search terms)
- ☐ Fuzzy (allows for typos and misspellings) Tolerance: 1

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



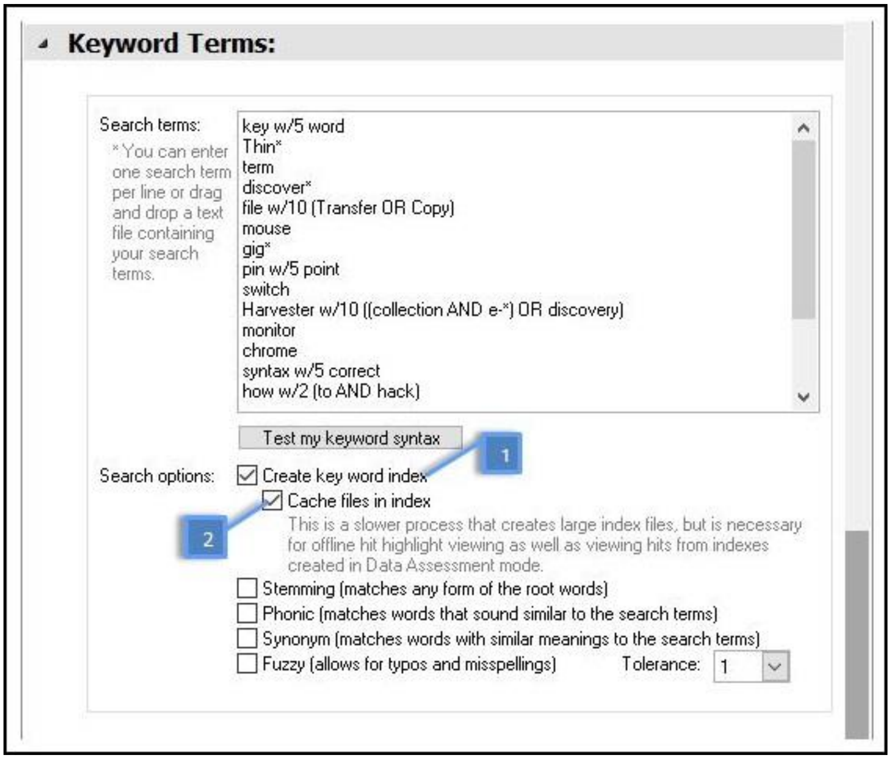
Keyword Terms:

Search terms: *You can enter one search term per line or drag and drop a text file containing your search terms.

```
key w/5 word
Thin*
term
discover*
file w/10 (Transfer OR Copy)
mouse
gig*
pin w/5 point
switch
Harvester w/10 ((collection AND e-*) OR discovery)
monitor
chrome
syntax w/5 correct
how w/2 (to AND hack)
```

Search options:
- ☑ Create key word index
  - ☑ Cache files in index
    This is a slower process that creates large index files, but is necessary for offline hit highlight viewing as well as viewing hits from indexes created in Data Assessment mode.
- ☐ Stemming (matches any form of the root words)
- ☐ Phonic (matches words that sound similar to the search terms)
- ☐ Synonym (matches words with similar meanings to the search terms)
- ☐ Fuzzy (allows for typos and misspellings) Tolerance: 1

NOTE: Creating full-text indexes **(1&2)** before the collection phase rather than from the collected information will create a search-speed reduction as Harvester creates indexes.



Keyword Terms:

Search terms:
*You can enter one search term per line or drag and drop a text file containing your search terms.

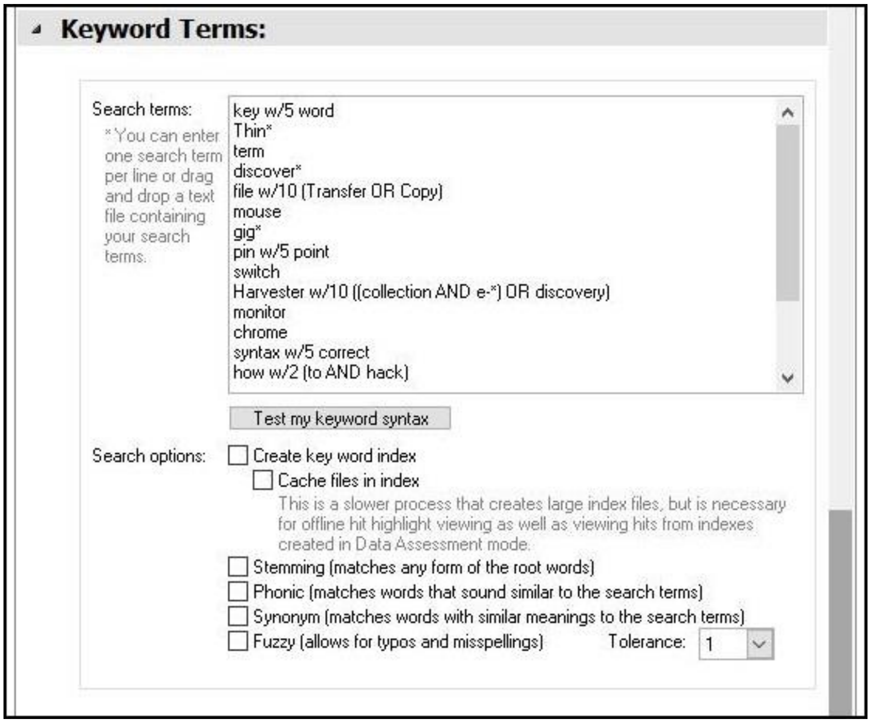
```
key w/5 word
Thin*
term
discover*
file w/10 (Transfer OR Copy)
mouse
gig*
pin w/5 point
switch
Harvester w/10 ((collection AND e-*) OR discovery)
monitor
chrome
syntax w/5 correct
how w/2 (to AND hack)
```

Test my keyword syntax

Search options:
☑ Create key word index
☑ Cache files in index
This is a slower process that creates large index files, but is necessary for offline hit highlight viewing as well as viewing hits from indexes created in Data Assessment mode.
☐ Stemming (matches any form of the root words)
☐ Phonic (matches words that sound similar to the search terms)
☐ Synonym (matches words with similar meanings to the search terms)
☐ Fuzzy (allows for typos and misspellings) Tolerance: 1

SEARCH TERM KEY WORD SYNTAX

Users can enter search terms and phrases as shown using one term per line. Harvester will treat the terms as ***OR***, flagging 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 outlined in the keyword syntax options.



Keyword Terms:

Search terms:
*You can enter one search term per line or drag and drop a text file containing your search terms.

```
key w/5 word
Thin*
term
discover*
file w/10 (Transfer OR Copy)
mouse
gig*
pin w/5 point
switch
Harvester w/10 ((collection AND e-*) OR discovery)
monitor
chrome
syntax w/5 correct
how w/2 (to AND hack)
```

Search options:
- □ Create key word index
  - □ Cache files in index
    This is a slower process that creates large index files, but is necessary for offline hit highlight viewing as well as viewing hits from indexes created in Data Assessment mode.
- □ Stemming (matches any form of the root words)
- □ Phonic (matches words that sound similar to the search terms)
- □ Synonym (matches words with similar meanings to the search terms)
- □ Fuzzy (allows for typos and misspellings) Tolerance: 1

Literal Search:

When searching for certain terms with two or more words, adding quotes around the term will prevent false keyword hits. Examples:

Management approval

would search the document for both words, in no specific order

"Management approval"

would search for these two words in specific order.

Document Keyword Search Supports Boolean Search Requests:

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)

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 custom 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 at the end of a word in the keyword 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%%%lphabet*** would hit on these as well as ***amphaket***.

<u>Proximity Searches:</u>

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

<u>Nested Searches</u>

(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**)

<u>Field Searches:</u>

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 are 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)**

<u>Automatic recognition of credit card numbers:</u>

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*)***

<u>Using keyword filters to extract from MBOX and DBX files:</u>

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***

<u>Test My Key Word Syntax:</u>

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

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.

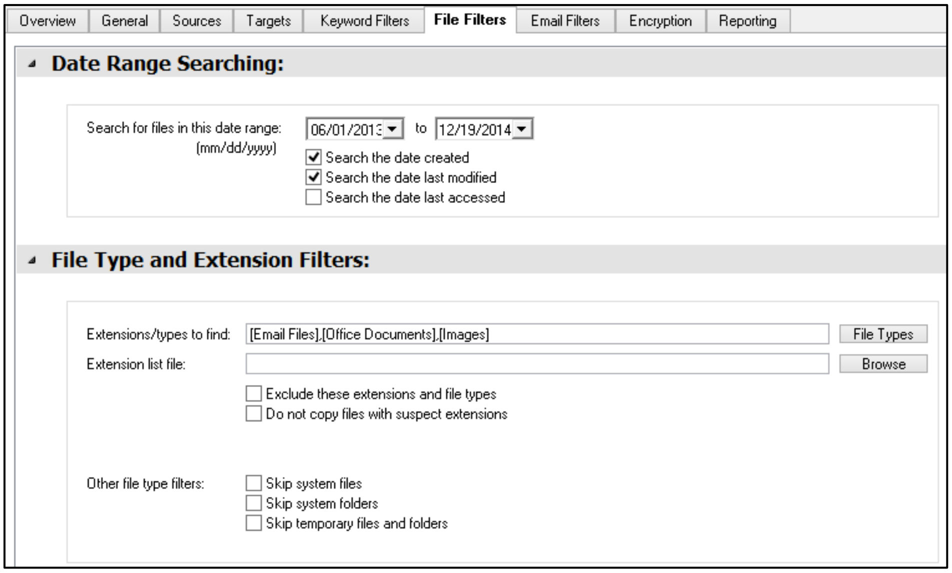
NOTE: If there is a fatal error in the key word syntax, Harvester will prompt you with the string(s) of syntax that are incorrect and warn you it will not be able to keyword search correctly with the error. For more information and how to fix the error, click the ***Test My Key Word Syntax*** button.

File Filters Tab

Harvester Portable 5.1 Advanced Options - Part 3

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.



Date Range Searching:

Search for files in this date range: (mm/dd/yyyy) 06/01/201: to 12/19/2014

☑ Search the date created
☑ Search the date last modified
☐ Search the date last accessed

File Type and Extension Filters:

Extensions/types to find: [Email Files],[Office Documents],[Images]

Extension list file:

☐ Exclude these extensions and file types
☐ Do not copy files with suspect extensions

Other file type filters:
☐ Skip system files
☐ Skip system folders
☐ Skip temporary files and folders

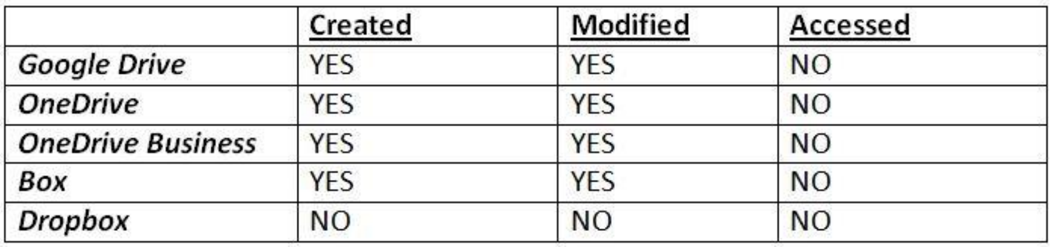
NOTE: One or more boxes must be checked for the date range to apply.

- Creation Dates
- Last Modified Dates
- Last Accessed Dates

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

<u>CLOUD SOURCE DATE RANGE</u>

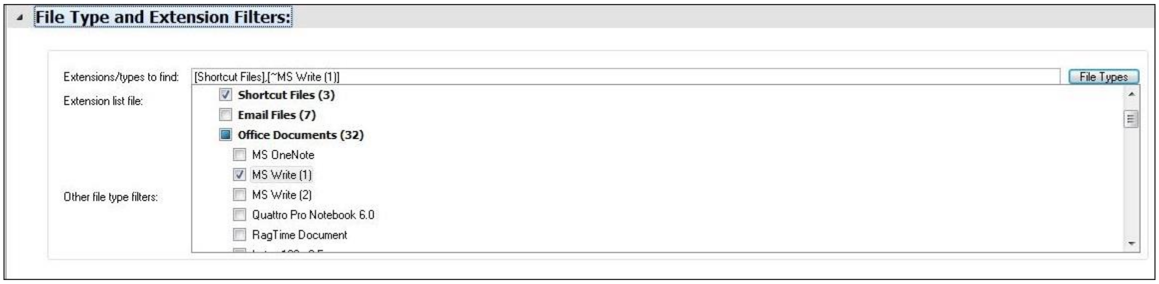
Not all dates are available for all cloud sources. Dates retrieved from cloud sources as created or modified will not necessarily be the same date as those displayed as created or modified via web browser. In some cases the displayed date will be the date that a file was uploaded, or a folder was created.



	<u>Created</u>	<u>Modified</u>	<u>Accessed</u>
Google Drive	YES	YES	NO
OneDrive	YES	YES	NO
OneDrive Business	YES	YES	NO
Box	YES	YES	NO
Dropbox	NO	NO	NO

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.



File Type and Extension Filters:

Extensions/types to find: [Shortcut Files],[~MS Write (1)]

Extension list file:

Other file type filters:

☑ **Shortcut Files (3)**
□ **Email Files (7)**
■ **Office Documents (32)**
□ MS OneNote
☑ MS Write (1)
□ MS Write (2)
□ Quattro Pro Notebook 6.0
□ RagTime Document

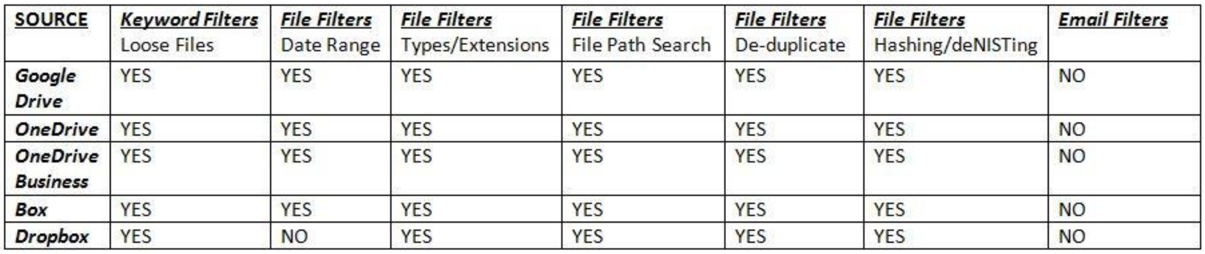
The following items can be specified:

- 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: Some file extensions have no header (TXT, INI, LOG, CSV), and therefore Harvester cannot perform a File Header Analysis. File Header Analysis is only performed when a file category is selected from the available options. Choosing file types for header signature filtering will likely 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.

CLOUD SOURCE FILTERS

Below is a table of filters and how they are applied to Cloud Sources:



SOURCE	*Keyword Filters* Loose Files	*File Filters* Date Range	*File Filters* Types/Extensions	*File Filters* File Path Search	*File Filters* De-duplicate	*File Filters* Hashing/deNISTing	*Email Filters*
Google Drive	YES	YES	YES	YES	YES	YES	NO
OneDrive	YES	YES	YES	YES	YES	YES	NO
OneDrive Business	YES	YES	YES	YES	YES	YES	NO
Box	YES	YES	YES	YES	YES	YES	NO
Dropbox	YES	NO	YES	YES	YES	YES	NO

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

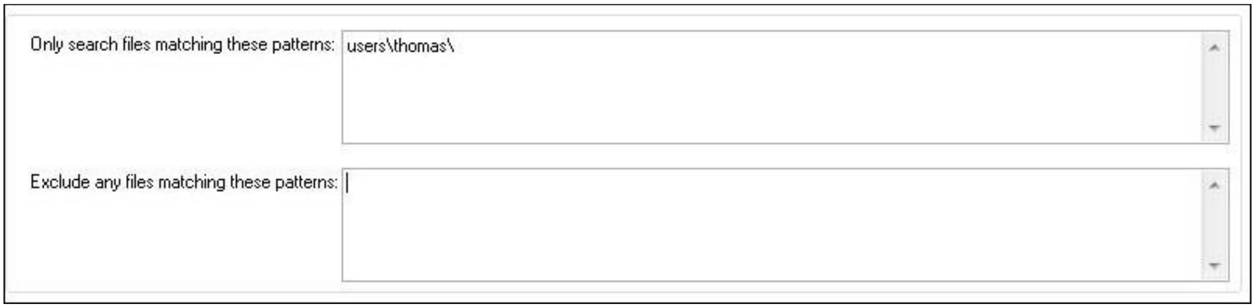
Exclude system files (with *System* attribute set): This option will filter out files which the file system (MFT/FAT) has flagged as system files. This is most commonly used in combination with the deNISTing option to further reduce the files collected.

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.

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. Example:



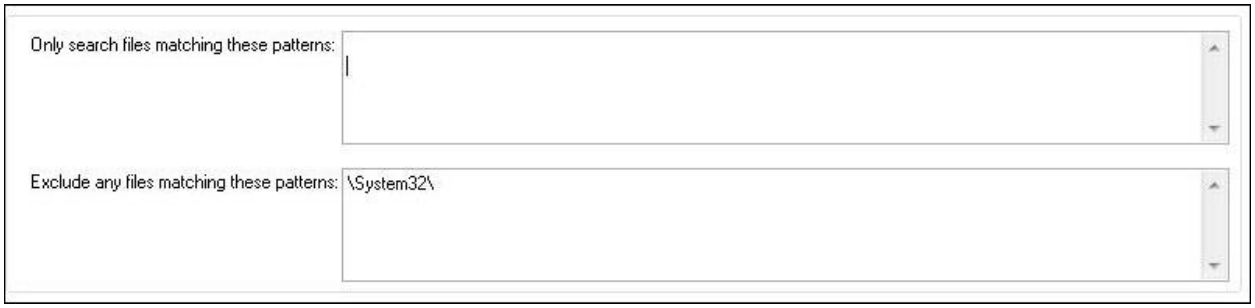
Only search files matching these patterns: users\thomas\

Exclude any files matching these patterns:

NOTE: These filters apply to whole paths. Comparisons are case insensitive

EXCLUDE FILES MATCHING THESE PATTERNS

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



Only search files matching these patterns:

Exclude any files matching these patterns: \System32\

Multiple patterns can be added if needed. The syntax options are listed on the main form.

Note: At time of comparison, all folders end with a \ character

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]).

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, or Gzip) or mail stores. An option to de-duplicate messages in email store 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 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

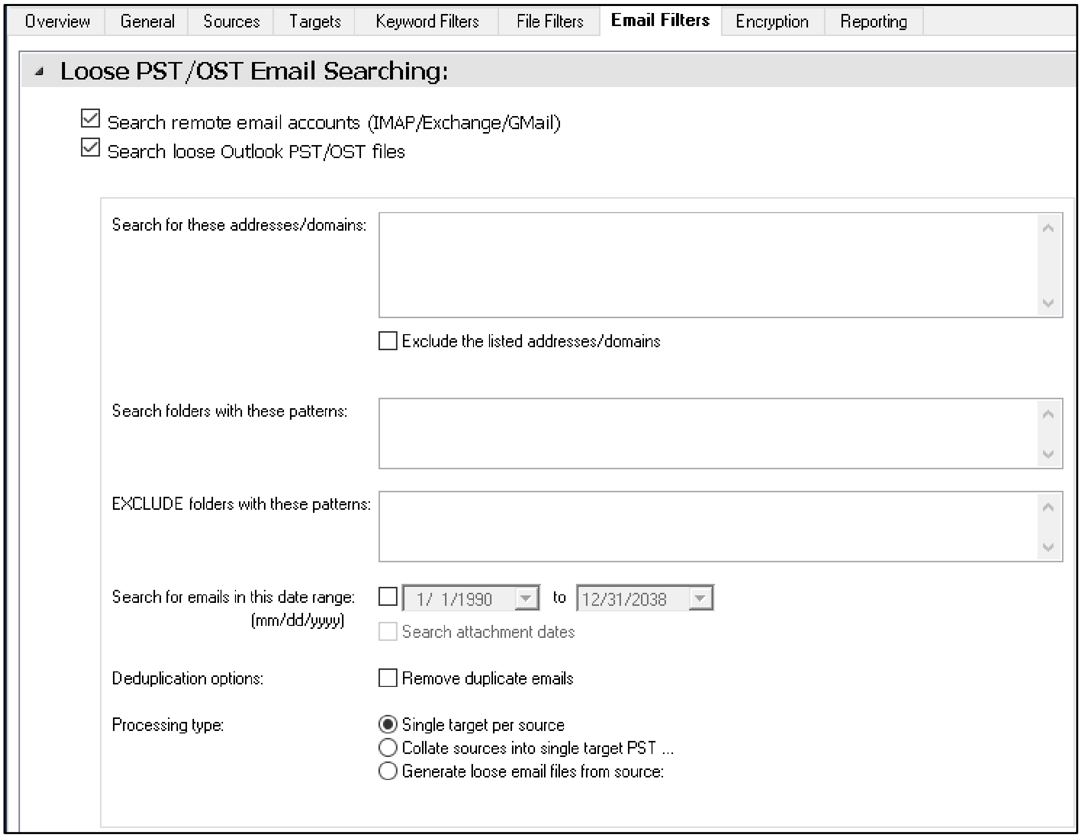
Email Filters Tab

Email Filters allow users to search individual Microsoft Outlook PST and Lotus Notes NSF files as well as *Active* email accounts including Microsoft Exchange.

Selecting the corresponding box expands the options for each section.

PST EMAIL SEARCHING – SEARCH LOOSE OUTLOOK PST FILES

Harvester 5.1 has the ability to search and copy messages from loose Outlook PST's and Exchange OST's without using Outlook or a MAPI connection. 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 is user customizable.



Loose PST/OST Email Searching:

☑ Search remote email accounts (IMAP/Exchange/GMail)
☑ Search loose Outlook PST/OST files

Search for these addresses/domains:

□ Exclude the listed addresses/domains

Search folders with these patterns:

EXCLUDE folders with these patterns:

Search for emails in this date range: (mm/dd/yyyy) □ 1/ 1/1990 to 12/31/2038
□ Search attachment dates

Deduplication options: □ Remove duplicate emails

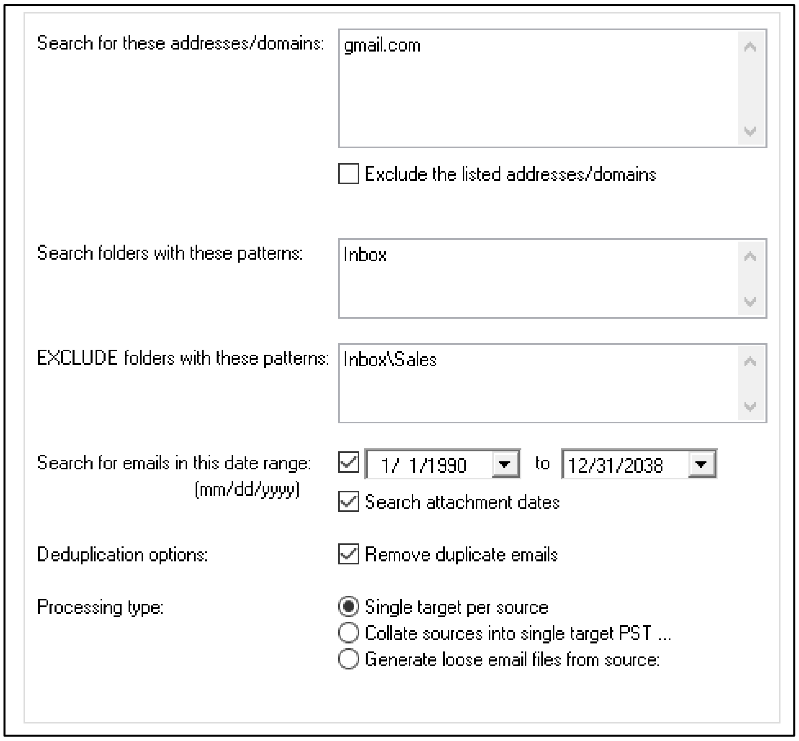
Processing type:
● Single target per source
○ Collate sources into single target PST ...
○ Generate loose email files from source:

Address/Domain to Search for (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 ***Address/Domain to Search For (To/From/CC/BCC)*** filter should be included in the results or excluded from them.

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. Leaving this field blank will search all folders.



Search for these addresses/domains: gmail.com

□ Exclude the listed addresses/domains

Search folders with these patterns: Inbox

EXCLUDE folders with these patterns: Inbox\Sales

Search for emails in this date range: (mm/dd/yyyy) ☑ 1/ 1/1990 to 12/31/2038

☑ Search attachment dates

Deduplication options: ☑ Remove duplicate emails

Processing type:
● Single target per source
○ Collate sources into single target PST ...
○ Generate loose email files from source:

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 ***InboxMyStuff/SKIP_THIS_FOLDER*** would be excluded.

These fields also support 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: Attachments to 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 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 ***_duplicate_emails.log*** file. 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.

Collate sources into a single target PST: This option will combine all source PSTs into the target PST 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 individual message files such as msg or eml.

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:

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

NOTE: OST files created with Outlook 2013 or newer cannot be searched as loose files with Harvester.

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

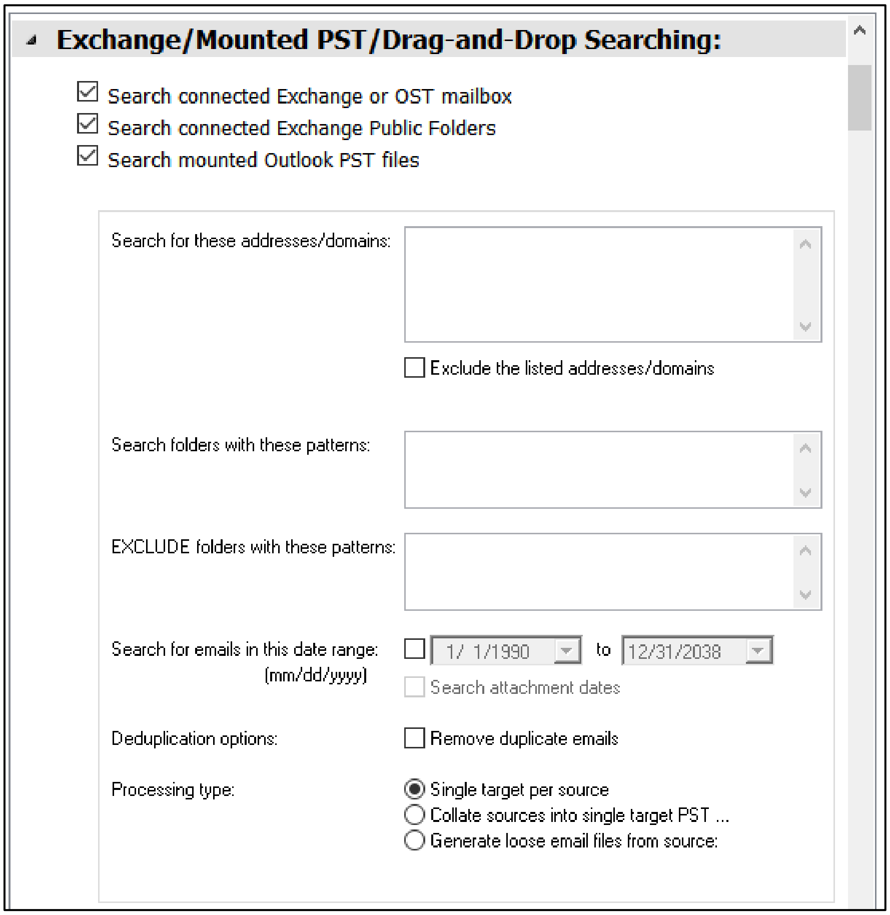
These options allow you to apply filtering to MS Exchange Mailboxes, PST files that are actively mounted in the user's Outlook, or Exchange Public Folders.

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

Search Connected Exchange Public Folders:
When checked, this searches and exports the resulting responsive emails from Exchange Public Folders.

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

Address/Domain to Search for (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.



Exchange/Mounted PST/Drag-and-Drop Searching:

☑ Search connected Exchange or OST mailbox
☑ Search connected Exchange Public Folders
☑ Search mounted Outlook PST files

Search for these addresses/domains:

☐ Exclude the listed addresses/domains

Search folders with these patterns:

EXCLUDE folders with these patterns:

Search for emails in this date range: (mm/dd/yyyy) ☐ 1/ 1/1990 to 12/31/2038
☐ Search attachment dates

Deduplication options: ☐ Remove duplicate emails

Processing type:
● Single target per source
○ Collate sources into single target PST ...
○ Generate loose email files from source:

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

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. 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 ***InboxMyStuff/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: Attachments to 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 stored emails 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 ***_duplicate_emails.log*** file. 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 original active.pst containing copies of the filtered messages. The new PST will reside in a path in the target according to the Target path in a folder called ***_ACTIVE_EMAIL_***.

Collate sources into a single target PST:
This option will combine all source PSTs into the target PST 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 individual message files such as msg or eml.

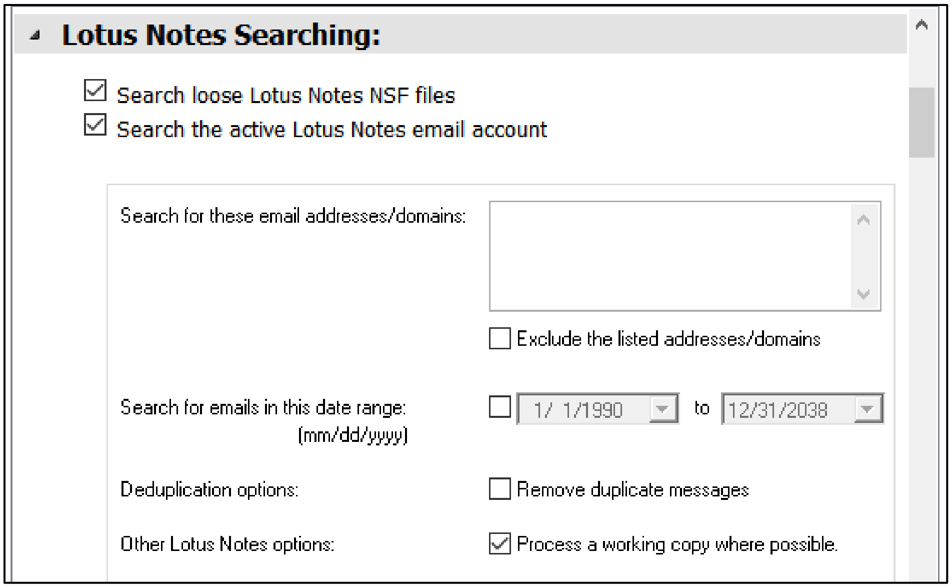
Export Type:
This option allows you to specify the format for the extracted messages. A copy of each email matching the filtered criteria will be saved in the chosen format and the subject is used as the filename. The messages will be stored in the same folder structure from the PST and the parent level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

- **ASCII 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 Lotus Notes:
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.



Lotus Notes Searching:

☑ Search loose Lotus Notes NSF files

☑ Search the active Lotus Notes email account

Search for these email addresses/domains:

□ Exclude the listed addresses/domains

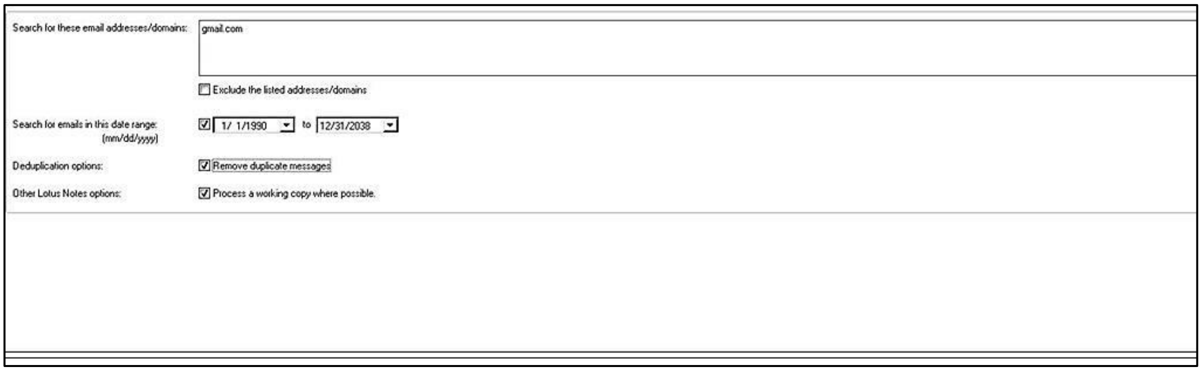
Search for emails in this date range: (mm/dd/yyyy) □ 1/ 1/1990 to 12/31/2038

Deduplication options: □ Remove duplicate messages

Other Lotus Notes options: ☑ Process a working copy where possible.

Search Active 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 Lotus Notes*** option.

Address/Domain to Search for (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.



Search for these email addresses/domains: gmail.com

□ Exclude the listed addresses/domains

Search for emails in this date range: (mm/dd/yyyy) ☑ 1/ 1/1990 to 12/31/2038

Deduplication options: ☑ Remove duplicate messages

Other Lotus Notes options: ☑ Process a working copy where possible.

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

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 ***_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

NOTE: *If you are using email filtering in any of the above options* ***and also 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)

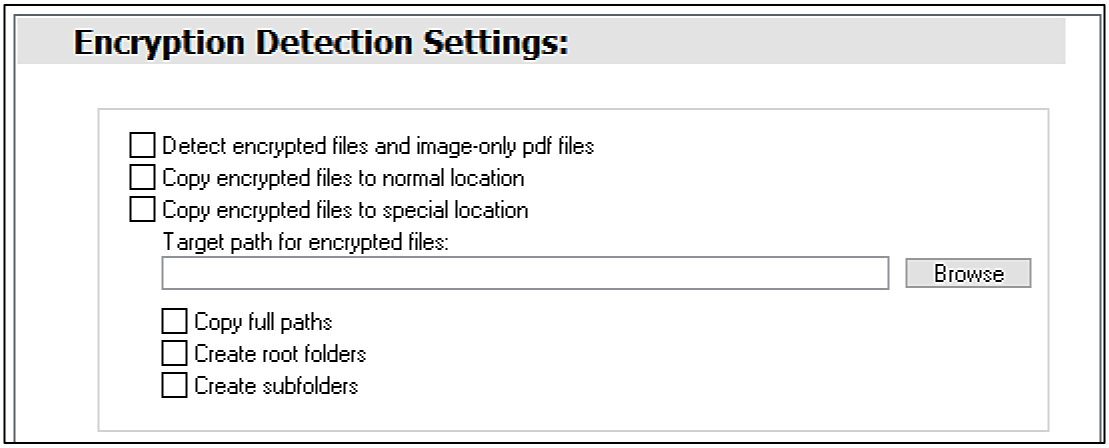
The PPLM email processing engine was introduced in Harvester 5.0 and is what is used to process messages from all mail sources except Lotus Notes. PPLM is multi-threaded and a sub folder will be created for each mail store that runs during a job.

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 Detection & Reporting Tabs

Harvester Portable 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 help determine whether to look for encrypted files and what to do with them if they are found.



Encryption Detection Settings:

□ Detect encrypted files and image-only pdf files

□ Copy encrypted files to normal location

□ Copy encrypted files to special location

Target path for encrypted files:

□ Copy full paths

□ Create root folders

□ Create subfolders

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.

Count Encrypted Files and Image-only files as KW hits:

Checking this box will count any encrypted document that also matches all other filter criteria (except keyword) as responsive. Unchecking this box will not flag encrypted items as responsive; it will only log them.

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 you to specify a target folder for encrypted files. You can either click the **Browse** button to select a folder or you can 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.

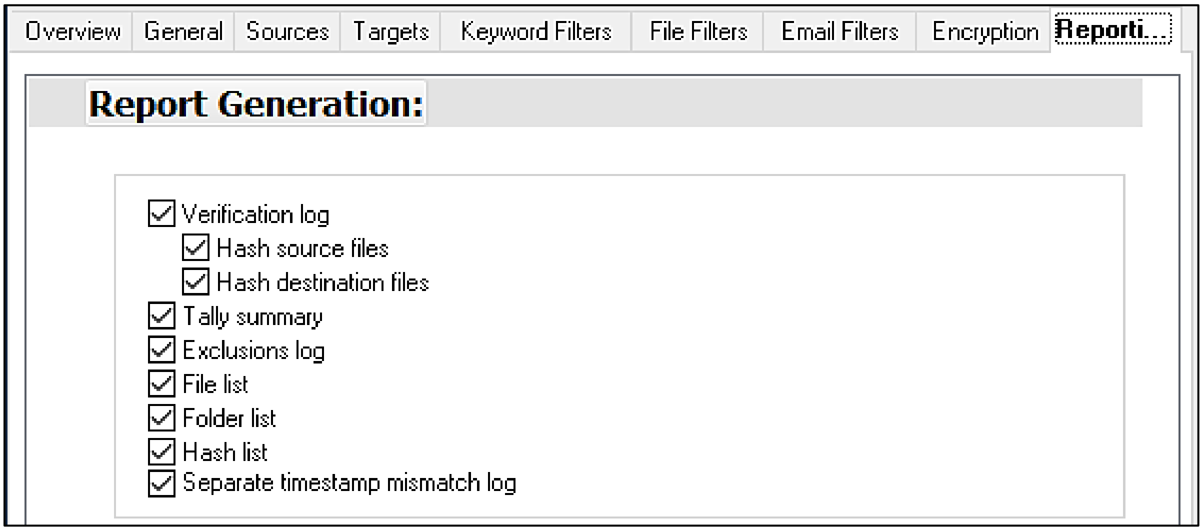
- **[Date]** – The date the job was run.
- **[DateTime]** – The date and time (to the second) that the job was run.

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:\demotest3\Crypto.doc* may be copied to *J:\Collected Files\Encrypted\C\demotest3\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



Report Generation:

- ☑ Verification log
  - ☑ Hash source files
  - ☑ Hash destination files
- ☑ Tally summary
- ☑ Exclusions log
- ☑ File list
- ☑ Folder list
- ☑ Hash list
- ☑ Separate timestamp mismatch log

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 files copied 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.

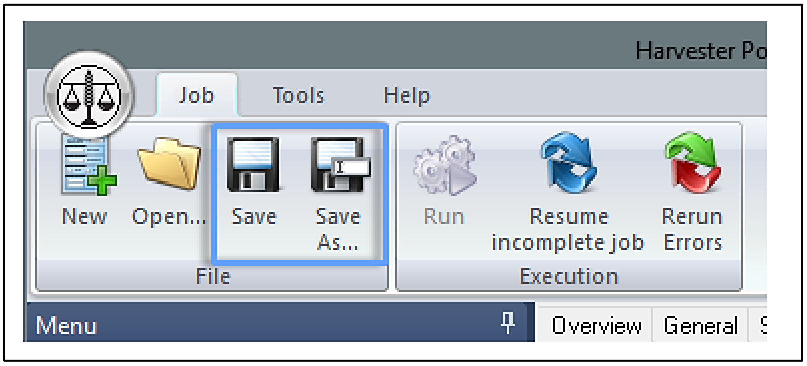
Folder List:
This option stores a file in the specified log path that contains the top-level folders specified as sources.

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 History

Reference: **Harvester Portable 5.1 Advanced Options Part 4**

Harvester provides quick access to previously run job statistics and settings through its History feature. A history database (History.db) is located in the Harvester application folder and will store the location and overall statistics of each completed job that is run from that Harvester executable. If you are running multiple Harvester Portable licenses, then each install will contain its own history database.

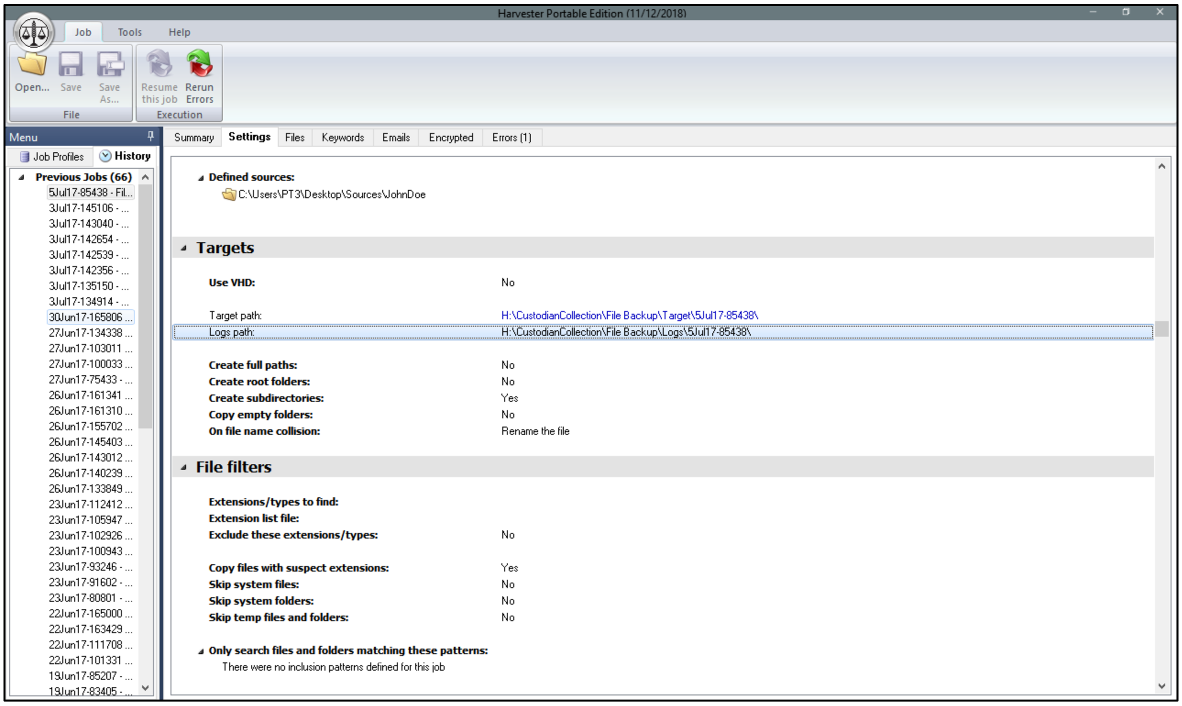
After a Harvester job completes, the progress bar will disappear and highlight the job history file that displays the ending job statistics and other useful details. If users create a new install from the Harvester archive file, then a new history database will be created. At times, there may be a need to access statistics from other locations. To access the job results from another location (not in the current list).

1. Click on **History** tab.
2. Click **Open** in the Job Ribbon.
3. Browse to the Harvester logs folder (contains .scj file) and click **Ok.**

Harvester will display the job in the History tab tree 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 that include 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.

To quickly access the Target Folder or Logs Folder, go to the Settings tab and double-click their corresponding path:



Defined sources:
C:\Users\PT3\Desktop\Sources\JohnDoe

Targets

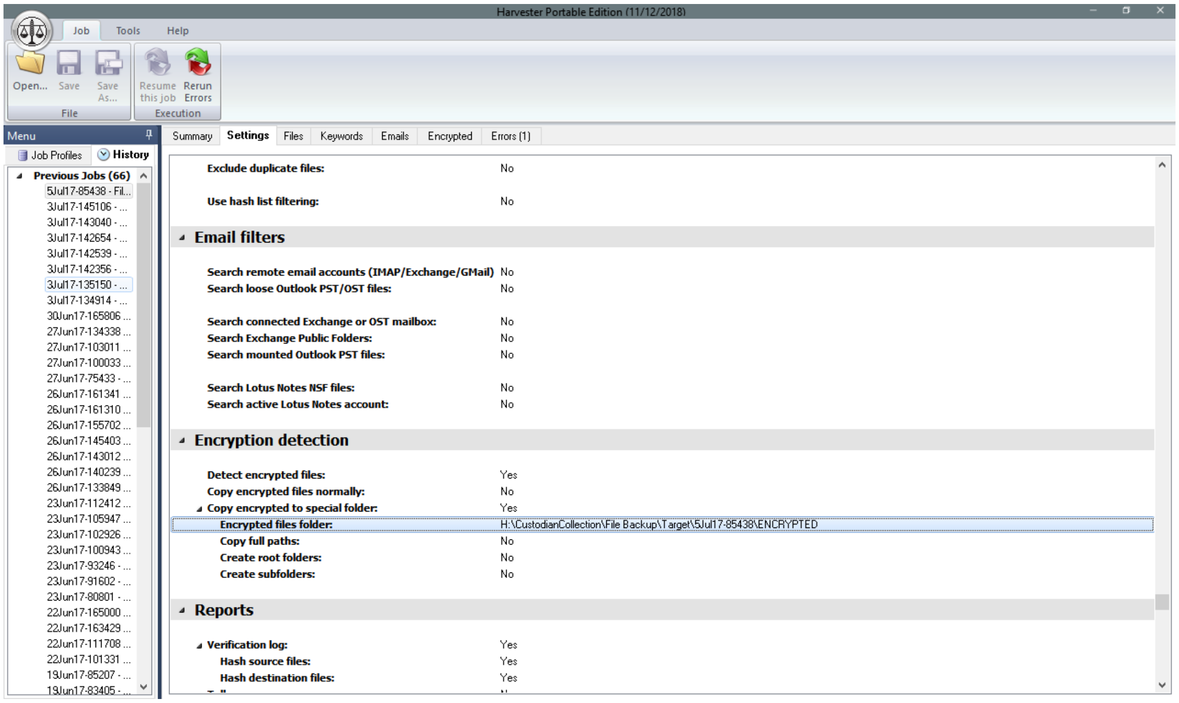
Use VHD:	No
Target path:	H:\CustodianCollection\File Backup\Target\5Jul17-85438\
Logs path:	H:\CustodianCollection\File Backup\Logs\5Jul17-85438\
Create full paths:	No
Create root folders:	No
Create subdirectories:	Yes
Copy empty folders:	No
On file name collision:	Rename the file

File filters

Extensions/types to find:	
Extension list file:	
Exclude these extensions/types:	No
Copy files with suspect extensions:	Yes
Skip system files:	No
Skip system folders:	No
Skip temp files and folders:	No

Only search files and folders matching these patterns:
There were no inclusion patterns defined for this job

This also applies to the user-added folders such as the emails and encrypted files paths



Exclude duplicate files:	No
Use hash list filtering:	No

Email filters

Search remote email accounts (IMAP/Exchange/GMail)	No
Search loose Outlook PST/OST files:	No
Search connected Exchange or OST mailbox:	No
Search Exchange Public Folders:	No
Search mounted Outlook PST files:	No
Search Lotus Notes NSF files:	No
Search active Lotus Notes account:	No

Encryption detection

Detect encrypted files:	Yes
Copy encrypted files normally:	No
Copy encrypted to special folder:	Yes
Encrypted files folder:	H:\CustodianCollection\File Backup\Target\5Jul17-85438\ENCRYPTED
Copy full paths:	No
Create root folders:	No
Create subfolders:	No

Reports

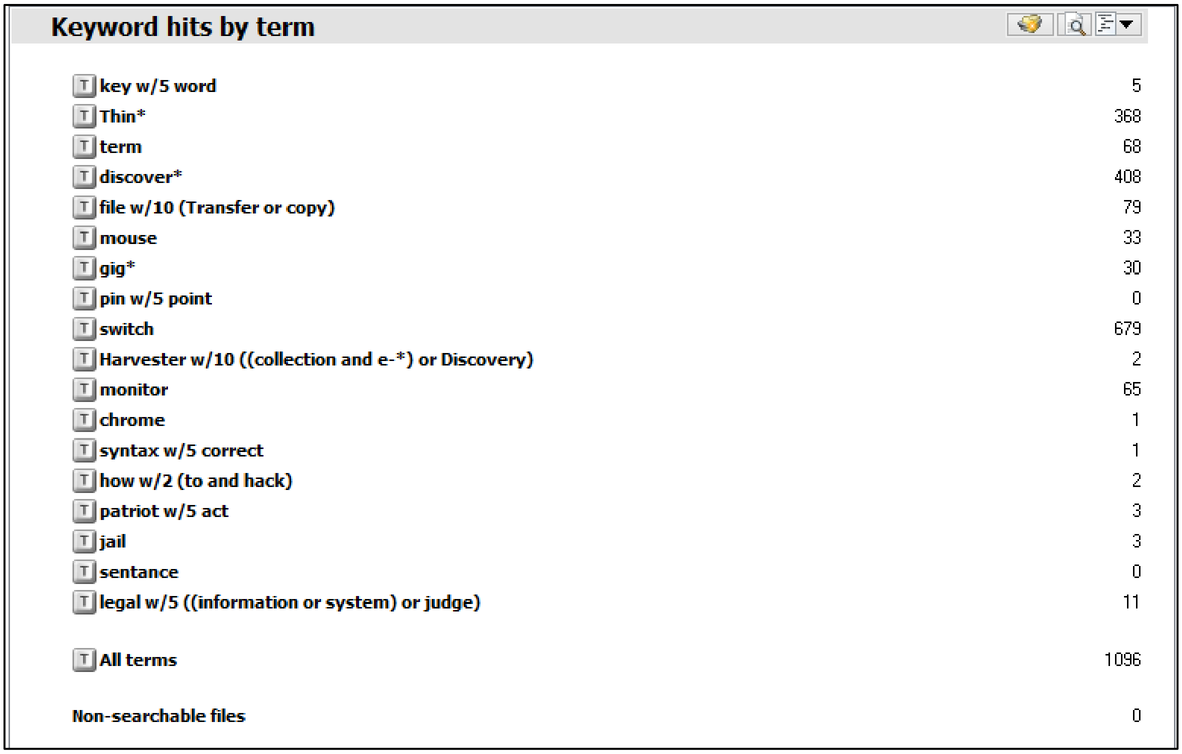
Verification log:	Yes
Hash source files:	Yes
Hash destination files:	Yes

KEYWORDS AND HIT HIGHLIGHT REVIEW

Reference: 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.

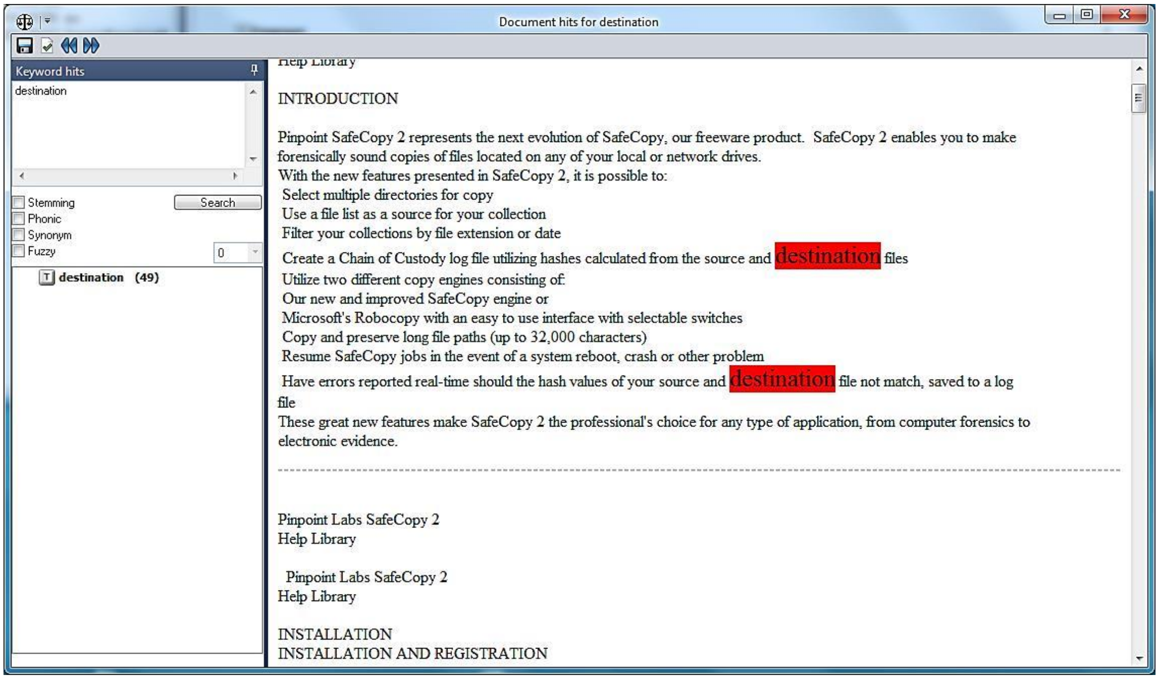
All entries will be listed in the ***Keywords*** tab as well at the total number of hits.



Keyword hits by term

key w/5 word	5
Thin*	368
term	68
discover*	408
file w/10 (Transfer or copy)	79
mouse	33
gig*	30
pin w/5 point	0
switch	679
Harvester w/10 ((collection and e-*) or Discovery)	2
monitor	65
chrome	1
syntax w/5 correct	1
how w/2 (to and hack)	2
patriot w/5 act	3
jail	3
sentance	0
legal w/5 ((information or system) or judge)	11
All terms	1096
Non-searchable files	0

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).



Help Library

INTRODUCTION

Pinpoint SafeCopy 2 represents the next evolution of SafeCopy, our freeware product. SafeCopy 2 enables you to make forensically sound copies of files located on any of your local or network drives.
With the new features presented in SafeCopy 2, it is possible to:
Select multiple directories for copy
Use a file list as a source for your collection
Filter your collections by file extension or date
Create a Chain of Custody log file utilizing hashes calculated from the source and destination files
Utilize two different copy engines consisting of:
Our new and improved SafeCopy engine or
Microsoft's Robocopy with an easy to use interface with selectable switches
Copy and preserve long file paths (up to 32,000 characters)
Resume SafeCopy jobs in the event of a system reboot, crash or other problem
Have errors reported real-time should the hash values of your source and destination file not match, saved to a log file
These great new features make SafeCopy 2 the professional's choice for any type of application, from computer forensics to electronic evidence.

Pinpoint Labs SafeCopy 2
Help Library

Pinpoint Labs SafeCopy 2
Help Library

INSTALLATION
INSTALLATION AND REGISTRATION

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 reports 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

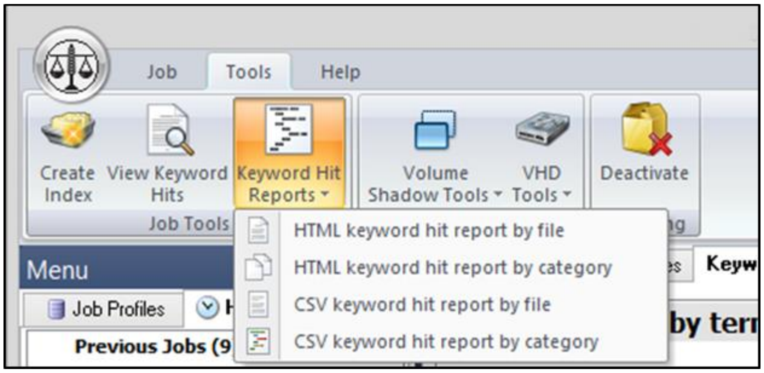
The keyword list panel in the lower left hand corner of the keyword preview window allows users to:

1. Click on an entry to preview the highlighted hits which are displayed in the right hand window
2. Tag files that are of interest to the current review process by clicking the checkbox next to each item.
3. Check multiple entries by using using ***CTRL+Shift (***select multiple entries) or holding Shift and clicking on1st then last entry to select a range. After selecting items right-click to ***Check selected items*** or ***Uncheck selected items***.
4. Remove files from the keyword list by right-clicking and selecting ***Remove checked items*** or ***Remove unchecked items***. Users can also click on the checkmark icon in the upper left hand corner.
5. 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

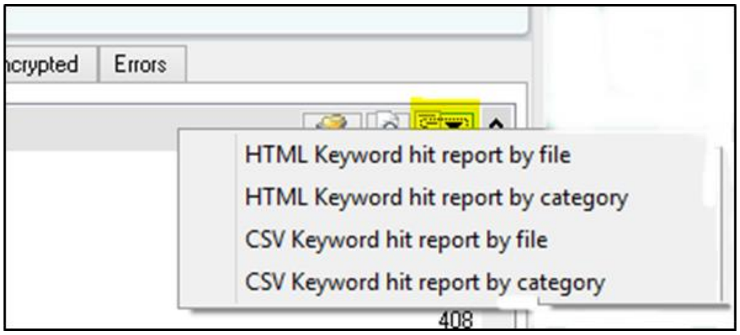
In addition to previewing the hit results, users can create HTML or CSV reports of all hits, individual entries or selected items. To create hit reports for all items:

1. Click on **History** tab.
2. Click the **Tools** menu.
3. Click on **Keyword Hit Reports**.
4. Select the report you want to generate and the location where you would like to store the file.



OR

1) Navigate to "**Keywords**" in "**History**".
2) Select "**Keyword Hit Reports**" located on the upper right side.
3) Select the option, then location where you would like to store the file.

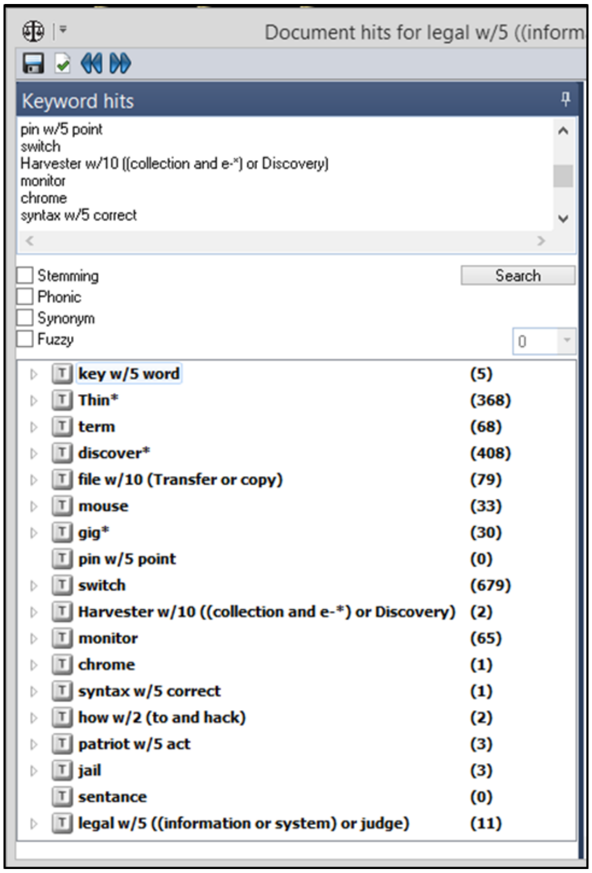

ncrypted
Errors
HTML Keyword hit report by file
HTML Keyword hit report by category
CSV Keyword hit report by file
CSV Keyword hit report by category
408

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

1. 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:
    - Click the check box next to each document.
    - Shift or Control keys select individual or a range of hits and right click will mark or unmark the highlighted items.
2. When finished tagging items click the icon to remove the remaining items from the list.
3. 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.



Document hits for legal w/5 ((inform

Keyword hits

```
pin w/5 point
switch
Harvester w/10 ((collection and e-*) or Discovery)
monitor
chrome
syntax w/5 correct
```

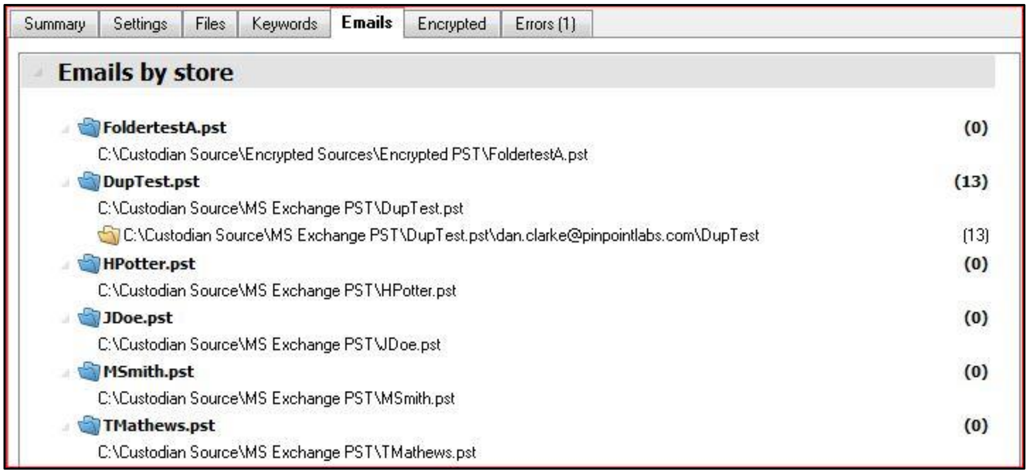
□ Stemming
□ Phonic
□ Synonym
□ Fuzzy

key w/5 word	(5)
Thin*	(368)
term	(68)
discover*	(408)
file w/10 (Transfer or copy)	(79)
mouse	(33)
gig*	(30)
pin w/5 point	(0)
switch	(679)
Harvester w/10 ((collection and e-*) or Discovery)	(2)
monitor	(65)
chrome	(1)
syntax w/5 correct	(1)
how w/2 (to and hack)	(2)
patriot w/5 act	(3)
jail	(3)
sentance	(0)
legal w/5 ((information or system) or judge)	(11)

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 **History Email** tab. Each mail store will be listed and allow users to expand to see the individual folders where the item are stored.

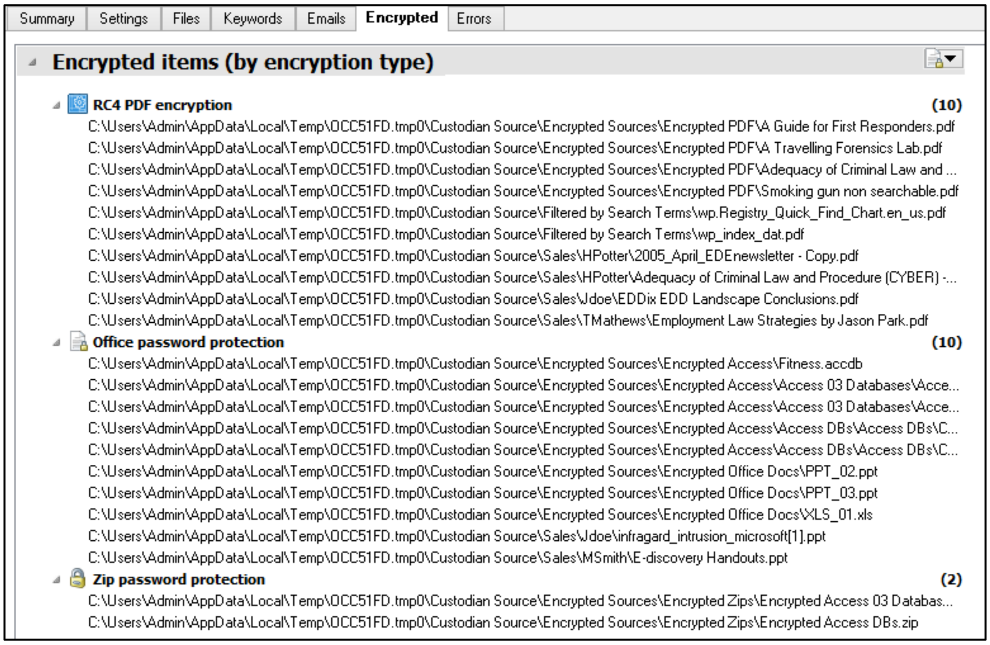


Emails by store

FoldertestA.pst (0)
C:\Custodian Source\Encrypted Sources\Encrypted PST\FoldertestA.pst
DupTest.pst (13)
C:\Custodian Source\MS Exchange PST\DupTest.pst
C:\Custodian Source\MS Exchange PST\DupTest.pst\dan.clarke@pinpointlabs.com\DupTest (13)
HPotter.pst (0)
C:\Custodian Source\MS Exchange PST\HPotter.pst
JDoe.pst (0)
C:\Custodian Source\MS Exchange PST\JDoe.pst
MSmith.pst (0)
C:\Custodian Source\MS Exchange PST\MSmith.pst
TMathews.pst (0)
C:\Custodian Source\MS Exchange PST\TMathews.pst

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 **History Encrypted Files** tab. Each category can be expanded to see the individual file locations.



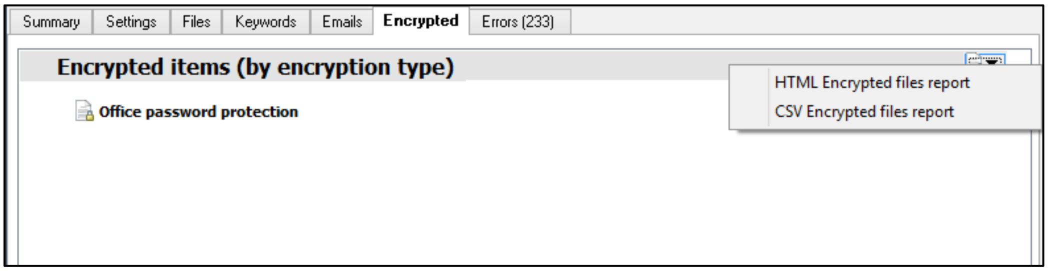
Encrypted items (by encryption type)

RC4 PDF encryption (10)
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted PDF\A Guide for First Responders.pdf
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted PDF\A Travelling Forensics Lab.pdf
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted PDF\Adequacy of Criminal Law and ...
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted PDF\Smoking gun non searchable.pdf
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Filtered by Search Terms\wp.Registry_Quick_Find_Chart.en_us.pdf
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Filtered by Search Terms\wp_index_dat.pdf
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Sales\HPotter\2005_April_EDEnewsletter - Copy.pdf
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Sales\HPotter\Adequacy of Criminal Law and Procedure (CYBER) -...
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Sales\Jdoe\EDDix EDD Landscape Conclusions.pdf
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Sales\TMathews\Employment Law Strategies by Jason Park.pdf
Office password protection (10)
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Access\Fitness.accdb
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Access\Access 03 Databases\Acce...
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Access\Access 03 Databases\Acce...
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Access\Access DBs\Access DBs\C...
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Access\Access DBs\Access DBs\C...
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Office Docs\PPT_02.ppt
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Office Docs\PPT_03.ppt
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Office Docs\XLS_01.xls
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Sales\Jdoe\infragard_intrusion_microsoft[1].ppt
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Sales\MSmith\E-discovery Handouts.ppt
Zip password protection (2)
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Zips\Encrypted Access 03 Databas...
C:\Users\Admin\AppData\Local\Temp\OCC51FD.tmp0\Custodian Source\Encrypted Sources\Encrypted Zips\Encrypted Access DBs.zip

In addition to previewing the encrypted items results, users can create HTML or CSV reports of encrypted items.
To create hit reports for encrypted items:

1) Click on "**History"** tab.
2) Click the "**Encrypted"** tab.
3) Click on "**Encrypted files Report**" icon on the right side of the window.

4) Select the option, then location where you would like to store the file.

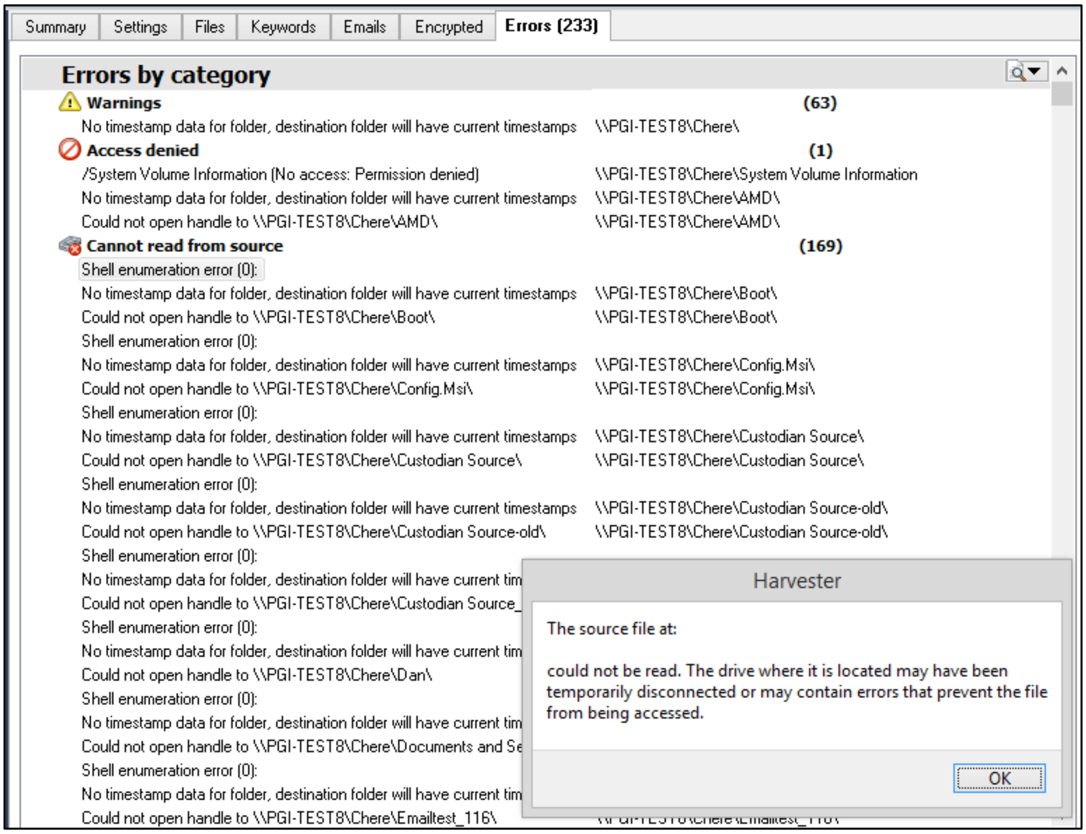


Encrypted items (by encryption type)

Office password protection

ERROR RESULTS

If errors occur during a Harvester job, they will be displayed in the **History 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.



Errors by category

```
Warnings                                                                                  (63)
   No timestamp data for folder, destination folder will have current timestamps   \\PGI-TEST8\Chere\
Access denied                                                                             (1)
   /System Volume Information (No access: Permission denied)                        \\PGI-TEST8\Chere\System Volume Information
   No timestamp data for folder, destination folder will have current timestamps   \\PGI-TEST8\Chere\AMD\
   Could not open handle to \\PGI-TEST8\Chere\AMD\                                  \\PGI-TEST8\Chere\AMD\
Cannot read from source                                                                   (169)
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current timestamps   \\PGI-TEST8\Chere\Boot\
   Could not open handle to \\PGI-TEST8\Chere\Boot\                                 \\PGI-TEST8\Chere\Boot\
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current timestamps   \\PGI-TEST8\Chere\Config.Msi\
   Could not open handle to \\PGI-TEST8\Chere\Config.Msi\                           \\PGI-TEST8\Chere\Config.Msi\
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current timestamps   \\PGI-TEST8\Chere\Custodian Source\
   Could not open handle to \\PGI-TEST8\Chere\Custodian Source\                     \\PGI-TEST8\Chere\Custodian Source\
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current timestamps   \\PGI-TEST8\Chere\Custodian Source-old\
   Could not open handle to \\PGI-TEST8\Chere\Custodian Source-old\                 \\PGI-TEST8\Chere\Custodian Source-old\
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current tim
   Could not open handle to \\PGI-TEST8\Chere\Custodian Source_
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current tim
   Could not open handle to \\PGI-TEST8\Chere\Dan\
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current tim
   Could not open handle to \\PGI-TEST8\Chere\Documents and Se
   Shell enumeration error (0):
   No timestamp data for folder, destination folder will have current tim
   Could not open handle to \\PGI-TEST8\Chere\Emailtest_116\
```

Harvester

The source file at:

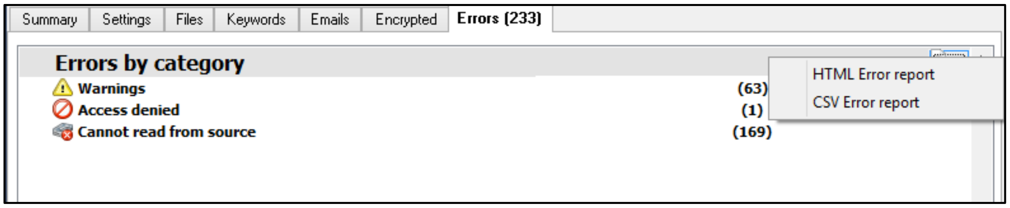
could not be read. The drive where it is located may have been temporarily disconnected or may contain errors that prevent the file from being accessed.

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***.

In addition to previewing the errors results, users can create HTML or CSV reports of encountered errors. To create error reports:

1) Click on "**History"** tab.
2) Click the "**Errors"** tab.
3) Click on "**Error report**" icon on the right side of the window.

Select the option, then location where you would like to store the file.



Errors by category

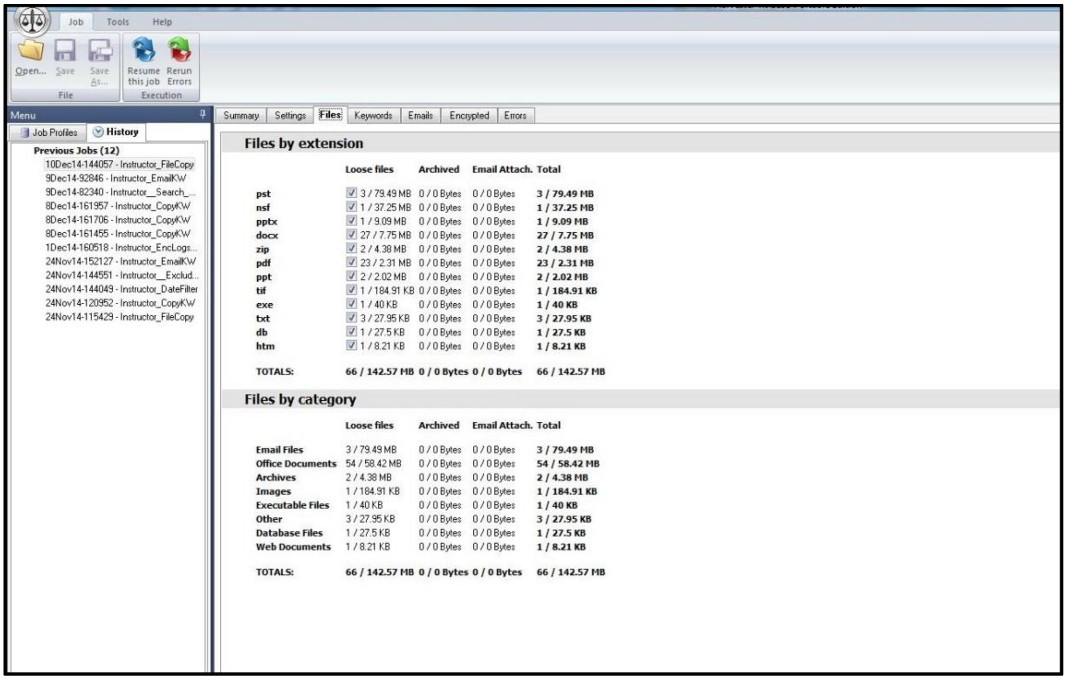
Warnings	(63)
Access denied	(1)
Cannot read from source	(169)

REMOVING UNWANTED FILE TYPES

Users have the option to see a file type summary and exclude file types from the final collection. To use this feature insure that ***Data Assessment*** from the ***General Tab*** is selected before running a job. **Data Assessment** mode will cause Harvester to stop before copying files and provide useful project summaries in the history tab as well as the logs folder.

To exclude unwanted file extensions follow these steps:

1. Click on the **History** tab.
2. Select the job from the list that you would like to review.
3. Click on the **Files** tab.
4. A checkbox will appear to the left of each extension that you would like to exclude.
5. Uncheck the box next to each extension that will not be copied.
6. Click **Save** in the **File** menu to save changes.
7. Click **Resume this job** from the **Execution** group to copy files with the selected extensions.



Previous Jobs (12)
10Dec14-144057 - Instructor_FileCopy
9Dec14-92846 - Instructor_EmailKW
9Dec14-82340 - Instructor__Search_...
8Dec14-161957 - Instructor_CopyKW
8Dec14-161706 - Instructor_CopyKW
8Dec14-161455 - Instructor_CopyKW
1Dec14-160518 - Instructor_EncLogs...
24Nov14-152127 - Instructor_EmailKW
24Nov14-144551 - Instructor__Exclud...
24Nov14-144049 - Instructor_DateFilter
24Nov14-120952 - Instructor_CopyKW
24Nov14-115429 - Instructor_FileCopy

Files by extension

	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

Files by category

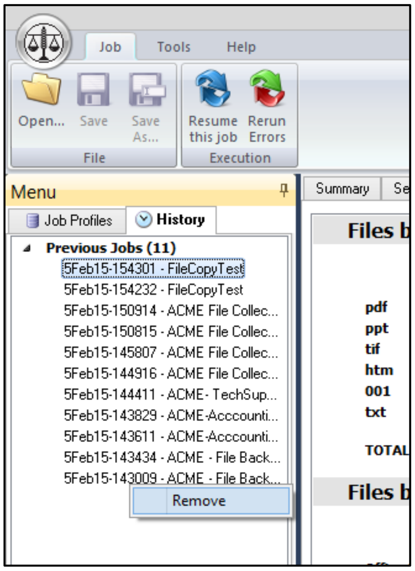
	Loose files	Archived	Email Attach.	Total
Email Files	3 / 79.49 MB	0 / 0 Bytes	0 / 0 Bytes	3 / 79.49 MB
Office Documents	54 / 58.42 MB	0 / 0 Bytes	0 / 0 Bytes	54 / 58.42 MB
Archives	2 / 4.38 MB	0 / 0 Bytes	0 / 0 Bytes	2 / 4.38 MB
Images	1 / 184.91 KB	0 / 0 Bytes	0 / 0 Bytes	1 / 184.91 KB
Executable Files	1 / 40 KB	0 / 0 Bytes	0 / 0 Bytes	1 / 40 KB
Other	3 / 27.95 KB	0 / 0 Bytes	0 / 0 Bytes	3 / 27.95 KB
Database Files	1 / 27.5 KB	0 / 0 Bytes	0 / 0 Bytes	1 / 27.5 KB
Web Documents	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

IMPORTANT NOTE: Once the job is resumed no further changes can be made to the extension list. Users will not be able to go back at a later time and copy files from the same job with different extensions.

When ***Save*** is selected the job database's excluded flag is updated for all records associated with the excluded file extensions. Harvester will resume the job copying files that match the selected extensions.

REMOVING UNWANTED HISTORY REPORTS

Users can remove old or unwanted history reports by right clicking the unwanted history report and selecting ***Remove***. Removing a history report does not delete any of the data associated with it. Target and log data can still be found in the path specified in the job file.

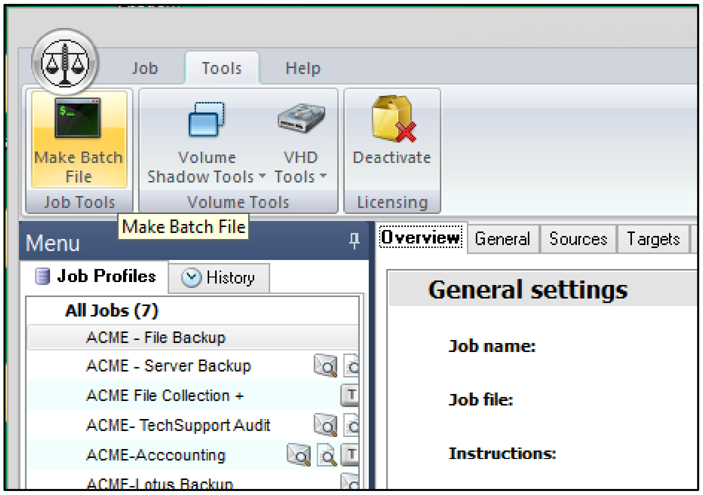


Previous Jobs (11)
5Feb15-154301 - FileCopyTest
5Feb15-154232 - FileCopyTest
5Feb15-150914 - ACME File Collec...
5Feb15-150815 - ACME File Collec...
5Feb15-145807 - ACME File Collec...
5Feb15-144916 - ACME File Collec...
5Feb15-144411 - ACME- TechSup...
5Feb15-143829 - ACME-Acccounti...
5Feb15-143611 - ACME-Acccounti...
5Feb15-143434 - ACME - File Back...
5Feb15-143009 - ACME - File Back...

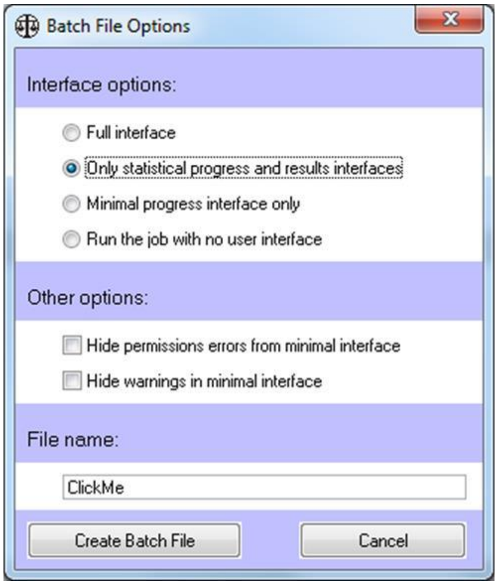
Tools Tab

MAKE A BATCH FILE

For self-collection kits, a critical tool is the option to create a batch file. If Harvester is running from a portable USB device, choose a job profile; click Tools and then Make Batch File.



The **Batch File Options** window will appear, offering interface options and how to display potential errors or warnings.



Interface options:

○ Full interface
● Only statistical progress and results interfaces
○ Minimal progress interface only
○ Run the job with no user interface

Other options:

☐ Hide permissions errors from minimal interface
☐ Hide warnings in minimal interface

File name:

ClickMe

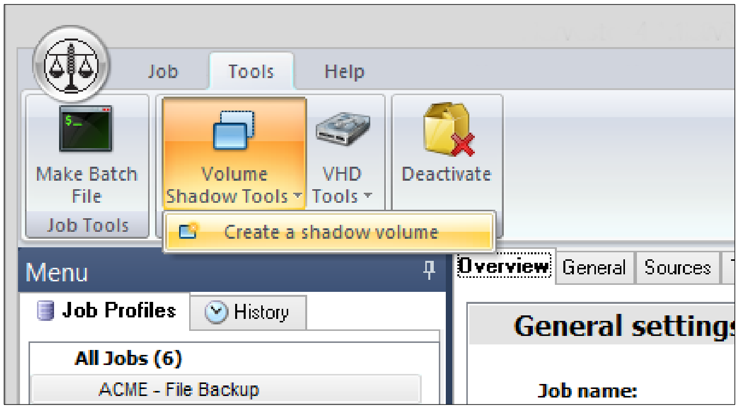
Harvester will create a ***clickme.bat*** in the root of the USB drive that Harvester is running from. In the batch file window, you have the option of choosing your own name for the batch file, rather than the default ***ClickMe***

If instructions or error notes were added in the Job Details, a Pop-up window will appear upon starting the *clickme.bat* unless ***Run the job with no user interface*** was chosen.

NOTE: The Harvester ESI Vault (which appears when ***Prompt user for additional sources*** is checked) is commonly used when distributing self-collection kits or jobs launched from Harvester for legal hold notice. By providing this interface in combination with the minimal progress interface, custodians can easily include items relevant to a matter without being notified of any of the other settings, sources, targets, or search criteria.
The ESI Vault interface is a drag-and-drop window that supports dragging and dropping relevant items.

VOLUME SHADOW TOOLS

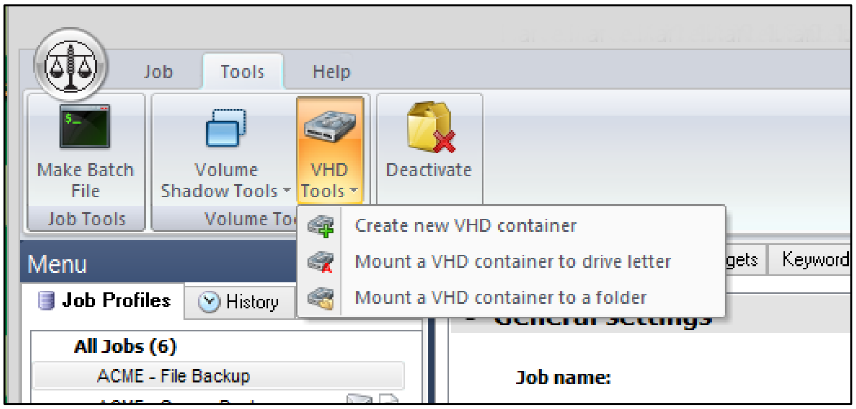
The Volume Shadow Tools can be found under the Tools tab in the upper ribbon.

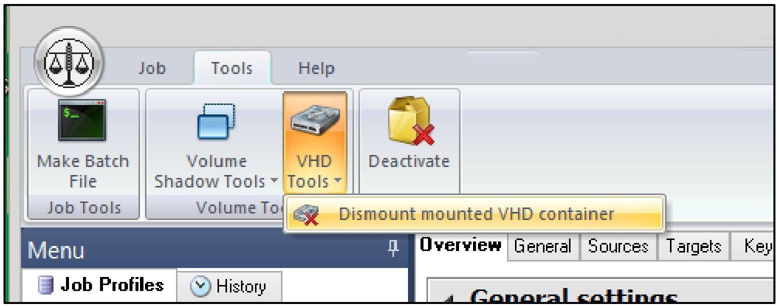


Harvester now utilizes **VSS**, or ***Volume Shadow Service***, to silently give access to files that are in-use or would have given permission errors by creating shadow volumes, read-only copies of files disks, so that files are able to be accessed without interfering with programs writing to those files. **The Volume Shadow Tools** can be found under the **Tools** tab in the upper ribbon. In order to create a volume shadow, you must be running as administrator and using Windows 7 or higher.

VIRTUAL HARD DRIVE TOOLS (VHD)

VHD Tools can be found under the **Tools Tab** in the upper ribbon





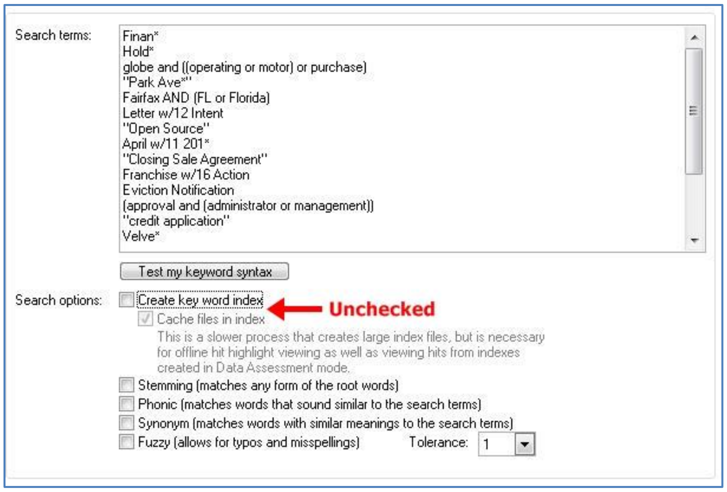
VHD (or Virtual Hard Drive) is a utility that creates file containers for collected data, keeping all collected data in a single container for easier transport. A VHD container file acts like any other kind of file, with the exception that it can also act as a hard drive on a Windows computer. Files that have been copied to this virtual hard drive will stay inside the VHD file. When you choose to copy your files to a VHD container, Harvester creates a VHD file at the location you 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.

You can use the VHD tools in the Tools ribbon to create a VHD container outside of a job or to mount an existing VHD container to view its contents. Presently, only VHD containers with single volumes are supported with this tool.
Note: In order to run VHD in Windows, Harvester must be run as Administrator.

Indexed vs Non-Indexed Searches

There are three different ways to go about a Harvester Keyword search:

1. **Non-Indexed -** Searching and copying data without indexing.



Search terms:

```
Finan*
Hold*
globe and ((operating or motor) or purchase)
"Park Ave*"
Fairfax AND (FL or Florida)
Letter w/12 Intent
"Open Source"
April w/11 201*
"Closing Sale Agreement"
Franchise w/16 Action
Eviction Notification
(approval and (administrator or management))
"credit application"
Velve*
```

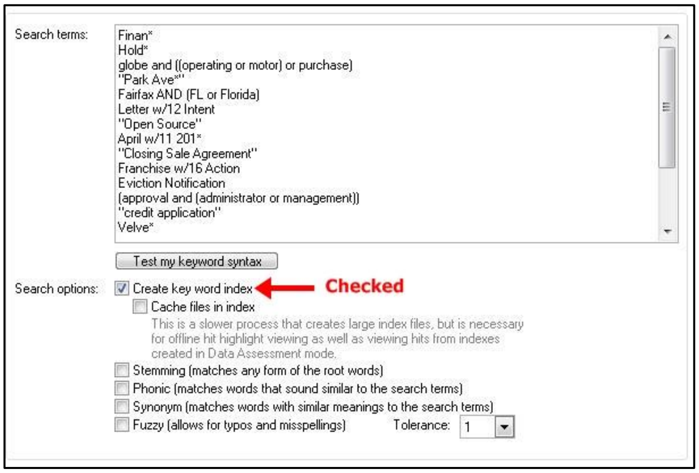
Search options:

- □ Create key word index ← Unchecked
  - ☑ Cache files in index
  - This is a slower process that creates large index files, but is necessary for offline hit highlight viewing as well as viewing hits from indexes created in Data Assessment mode.
- □ Stemming (matches any form of the root words)
- □ Phonic (matches words that sound similar to the search terms)
- □ Synonym (matches words with similar meanings to the search terms)
- □ Fuzzy (allows for typos and misspellings) Tolerance: 1

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.

2. <u>Create Keyword Index</u> - Search data and create a keyword index *of the copied data*.

This is most helpful with viewing keyword hit highlights



Search terms:

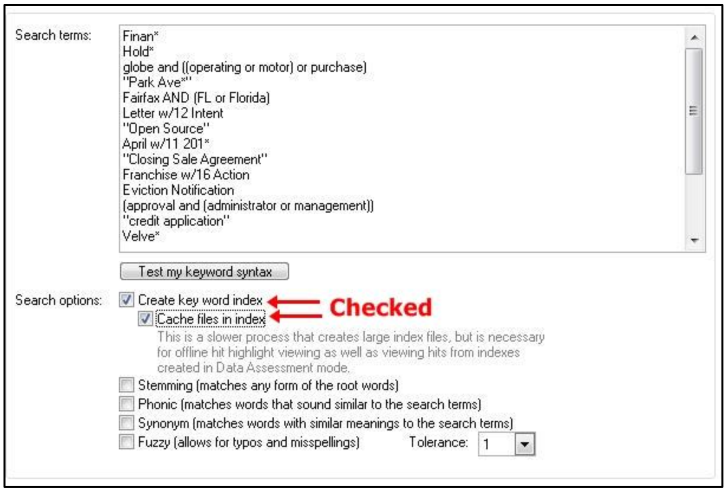
```
Finan*
Hold*
globe and ((operating or motor) or purchase)
"Park Ave*"
Fairfax AND (FL or Florida)
Letter w/12 Intent
"Open Source"
April w/11 201*
"Closing Sale Agreement"
Franchise w/16 Action
Eviction Notification
(approval and (administrator or management))
"credit application"
Velve*
```

Search options:

- ☑ Create key word index ← Checked
  - □ Cache files in index
  - This is a slower process that creates large index files, but is necessary for offline hit highlight viewing as well as viewing hits from indexes created in Data Assessment mode.
- □ Stemming (matches any form of the root words)
- □ Phonic (matches words that sound similar to the search terms)
- □ Synonym (matches words with similar meanings to the search terms)
- □ Fuzzy (allows for typos and misspellings) Tolerance: 1

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

3. <u>Cache File in Index</u> - Copies the entire contents of the files into the index.



Search terms:

```
Finan*
Hold*
globe and ((operating or motor) or purchase)
"Park Ave*"
Fairfax AND (FL or Florida)
Letter w/12 Intent
"Open Source"
April w/11 201*
"Closing Sale Agreement"
Franchise w/16 Action
Eviction Notification
(approval and (administrator or management))
"credit application"
Velve*
```

Search options:

- ☑ Create key word index
  - ☑ Cache files in index

    This is a slower process that creates large index files, but is necessary for offline hit highlight viewing as well as viewing hits from indexes created in Data Assessment mode.
- ☐ Stemming (matches any form of the root words)
- ☐ Phonic (matches words that sound similar to the search terms)
- ☐ Synonym (matches words with similar meanings to the search terms)
- ☐ Fuzzy (allows for typos and misspellings) Tolerance: 1

This can be applied to jobs run in Data Assessment Mode for situations where indexing needs to implement before (or without) copying the data.

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

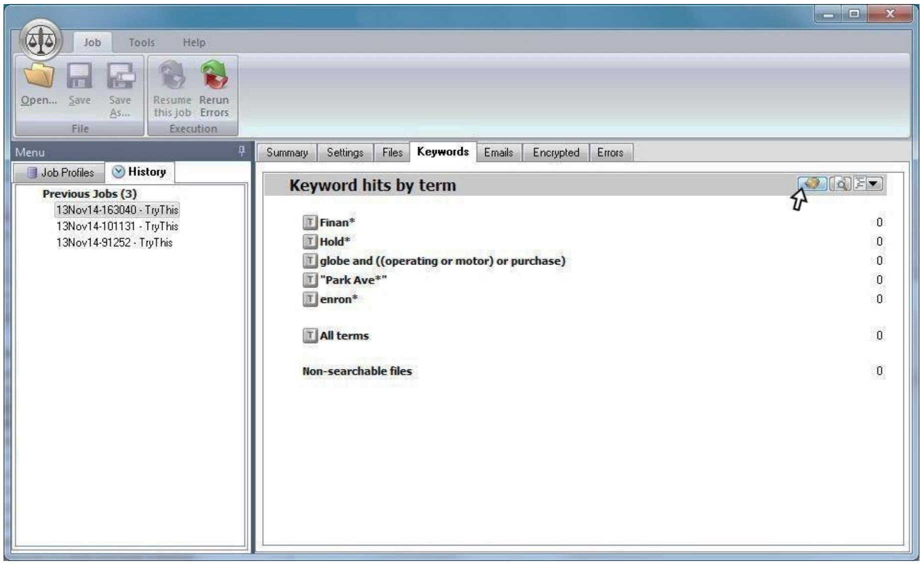
- Review keyword hit totals
- Create keyword hit reports
- View keyword hit highlighting for copied documents and email

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.

Users also have the option to create an index after copying is completed in the History > Keywords tab

- **Click the Create Index icon:**



Previous Jobs (3)
13Nov14-163040 - TryThis
13Nov14-101131 - TryThis
13Nov14-91252 - TryThis

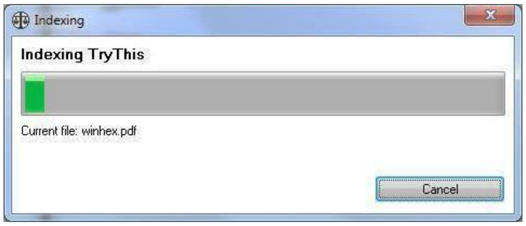
```
Keyword hits by term

Finan*                                          0
Hold*                                           0
globe and ((operating or motor) or purchase)    0
"Park Ave*"                                     0
enron*                                          0

All terms                                       0

Non-searchable files                            0
```

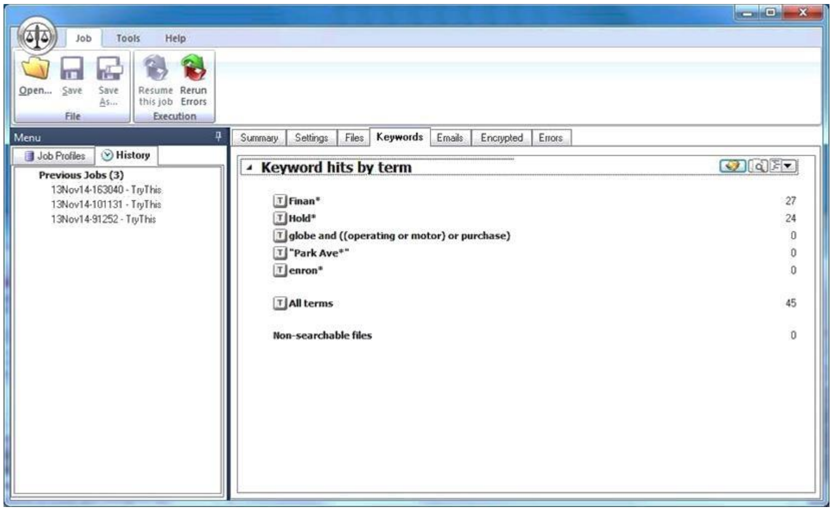
- **Indexing begins:**



Indexing TryThis

Current file: winhex.pdf

- **Indexing complete:**



Previous Jobs (3)
13Nov14-163040 - TryThis
13Nov14-101131 - TryThis
13Nov14-91252 - TryThis

```
Keyword hits by term

Finan*                                          27
Hold*                                           24
globe and ((operating or motor) or purchase)    0
"Park Ave*"                                     0
enron*                                          0

All terms                                       45

Non-searchable files                            0
```

.OCC/.SCJ File Structure Definitions

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]    JobName
[JINSTRUCT]       These Are Instructions
[JERRINST] Read This Message
[EDAPAUSE]        1
[THREADS]  0
[SUSPECTNOCOPY]   1
[WRITE_TO_VHD]    0
[VHD_PATH]        C:\Users\User\Desktop\VHD.vhd
[VHD_TARGET]      Target
[VHD_LOGS]        C:\Users\User\Desktop\Logs
[VHD_MOUNTPOINT]
[VHD_VOL_LABEL]
[WRITE_TO_VCC]    1
[VCCP]     7BD4D7ABD7812CC90214A6AFEEDA1CDC
[VCCCONTAINER]    C:\Users\User\Desktop\VC.vc
[VCCMOUNTLETTER]
[VCCTARGET]       Target
[VCCLOGS]  C:\Users\User\Desktop\Logs
[TARGET]   C:\Users\User\Desktop\Target
[JPATH]    C:\Users\User\Desktop\Logs
[COMMONLOG]
[DRIVETOUNC]      1
[SRC]      C:\Users\User\Desktop\Source
[SRC]      [IMAP=email@emailaddress.com|***********|imap.account.com|993|1]
[SRC]      [EXCH=email@emailaddress.com|***********|exchange.ews.url||]
[SRC]      GOOGLEDRIVE
[SRC]      ONEDRIVEBUSINESS
[SRC]      ONEDRIVE
[SRC]      DROPBOX
[SRC]      BOX
[PROMPT]   0
[USEVSS]   1
[ENGINE]   SC
[COPYEMPTIES]     1
[COPYSUBS]        1
[ZIPDIRS]  1
[FILTERESIV]      1
[COPYFILES]       1
[TALLYSUM]        1
[FILELIST] 1
[FOLDERLIST]      1
[BMONTH]   1
[BDAY]     1
[BYEAR]    1990
[EMONTH]   12
[EDAY]     31
[EYEAR]    2038
[SEARCHCREATED]   1
[SEARCHMODIFIED]  1
[SEARCHACCESSED]  1
[SILENT]   1
[OVERWRITE]       0
[RENAME]   1
[FULLPATH]        1
[ROOTFOLDERS]     1
[CREATESUBS]      1
```

```
[SEPARATE_TS]   1
[EXTOP] 0
[LOGEXCLUSIONS] 1
[LOG]   1
[S-HASH]        1
[D-HASH]        1
[SHOWSHELL]     1
[STARTEX]       qwerty
[SHELLEX]       poiuyt
[SEARCHEMAILS]  1
[SEARCHLIVEPST] 1
[SEARCHPUBLICFOLDERS]   1
[SEARCHEXCHANGEBOX]     1
[SEARCHLOTUS]   1
[KWSUBJECTBODY] 1
[KWEMAILHEADERS]        1
[ACTIVELOTUS]   1
[KWATTACH]      1
[KWCREATEINDEX] 1
[CACHEINDEX]    1
[KWSTEMMING]    1
[KWPHONIC]      1
[KWSYNONYM]     1
[KWFUZZY]       1
[KWFUZZYTOL]    1
[KWARCHIVEOPT]  0
[HASHFILTER]    1
[HFILTER_EMAILS]        1
[HFILTER_ATTACH]        1
[HASHFILTERINCLUDE]     0
[HASHFILTEREXCLUDE]     1
[KWGO]  1
[KWONLY] 1
[KWHITENCRYPTED]        1
[KWINCNSATTACH] 1
[DETECTENC]     1
[COPYENCTO]     1
[ENCTARGET]     [Target]\ENCRYPTED
[ENCFULLPATHS]  1
[ENCROOTFOLDERS]        1
[ENCSUBFOLDERS] 1
[COPYENCNORMAL] 1
[SKIPSYSFILES]  1
[SKIPSYSDIRS]   1
[SKIPTEMP]      1
[SKIPDUPES]     1
[HASHLIST]      1
[KWLIST]        term1
[KWLIST]        term2
[KWLIST]        term3
[EDATESEARCH]   0
[ESTARTMM]      1
[AEDATESEARCH]  0
[AESTARTMM]     1
[LNDATESEARCH]  0
[LNSTARTMM]     1
[ESTARTDD]      1
[AESTARTDD]     1
[LNSTARTDD]     1
[ESTARTYYYY]    1990
[AESTARTYYYY]   1990
[LNSTARTYYYY]   1990
[EENDMM]        12
[AEENDMM]       12
[LNENDMM]       12
[EENDDD]        31
```

```
[AEENDDD]          31
[LNENDDD]          31
[EENDYYYY]         2038
[AEENDYYYY]        2038
[LNENDYYYY]        2038
[EDEDUPE]0
[AEDEDUPE]         0
[EATTACHDATES]     0
[AEATTACHDATES]    0
[LNDEDUPE]         0
[LNWORKINGCOPY]    1
[EADDYEX]0
[KWEXCLUDE]        1
[AEADDYEX]         0
[LNADDYEX]         0
[LNLOGINTYPE]      0
[LNSERVER]
[LNUSER]
[LNPASS]
[EPTYPE]   0
[AEPTYPE]  0
[EEXPORTFORMAT]    0
[AEEXPORTFORMAT]   0
```

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.
[VHD_VOL_LABEL]	text	
[WRITE_TO_VCC]	1 or 0	
[VCCP]	text	
[VCCCONTAINER]	text	
VCCMOUNTLETTER]	text	
[VCCTARGET]	text	
[VCCLOGS]	text	
[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.
[USEVSS]	1 or 0	
[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 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.
[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.

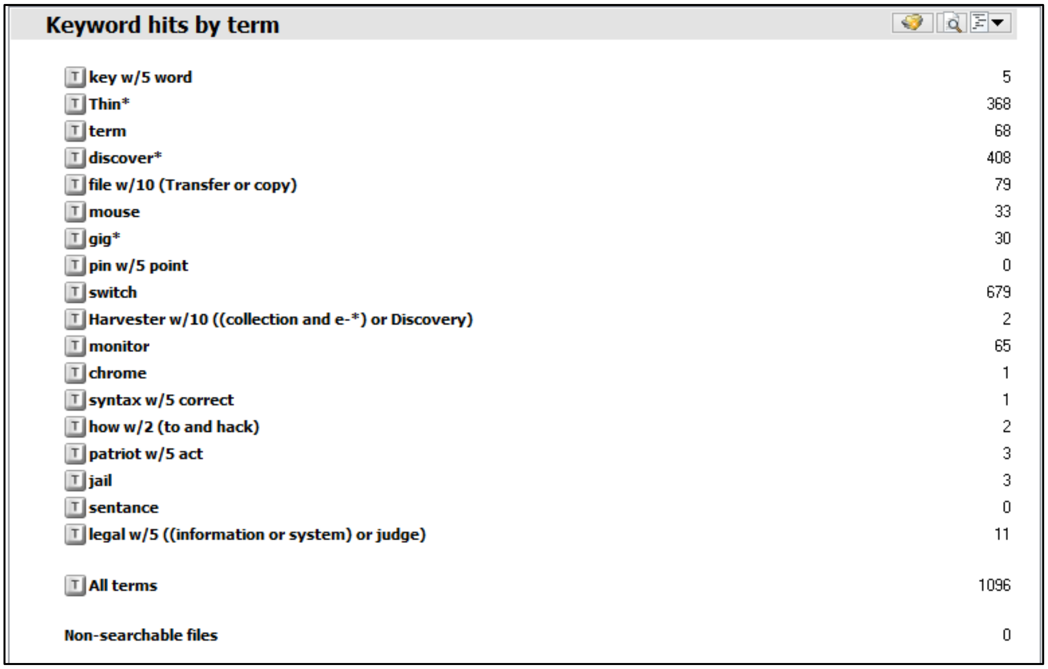
Keyword Highlighting and Indexes

KEYWORDS AND HIT HIGHLIGHT REVIEW

Reference: 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.

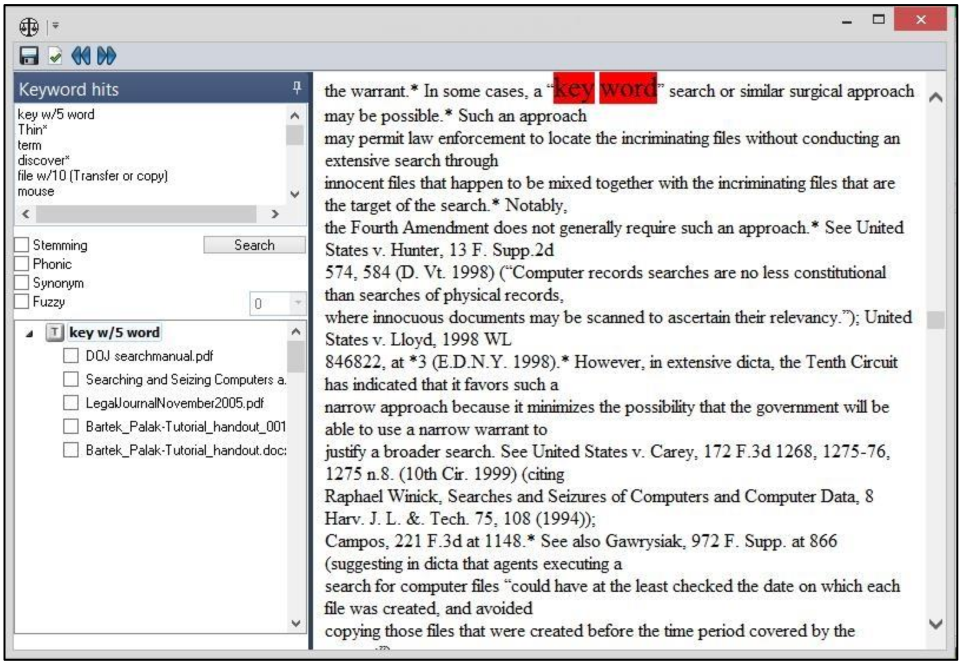
All entries will be listed in the ***Keywords*** tab as well at the total number of hits.



Keyword hits by term

key w/5 word	5
Thin*	368
term	68
discover*	408
file w/10 (Transfer or copy)	79
mouse	33
gig*	30
pin w/5 point	0
switch	679
Harvester w/10 ((collection and e-*) or Discovery)	2
monitor	65
chrome	1
syntax w/5 correct	1
how w/2 (to and hack)	2
patriot w/5 act	3
jail	3
sentance	0
legal w/5 ((information or system) or judge)	11
All terms	1096
Non-searchable files	0

Double clicking on an entry will bring up a window that lists the individual files in the left pane that have matching hits. Clicking on a file in the list will display the contents in the preview windows and matching terms will be highlighted (as seen below).



Keyword hits

key w/5 word
Thin*
term
discover*
file w/10 (Transfer or copy)
mouse

□ Stemming
□ Phonic
□ Synonym
□ Fuzzy

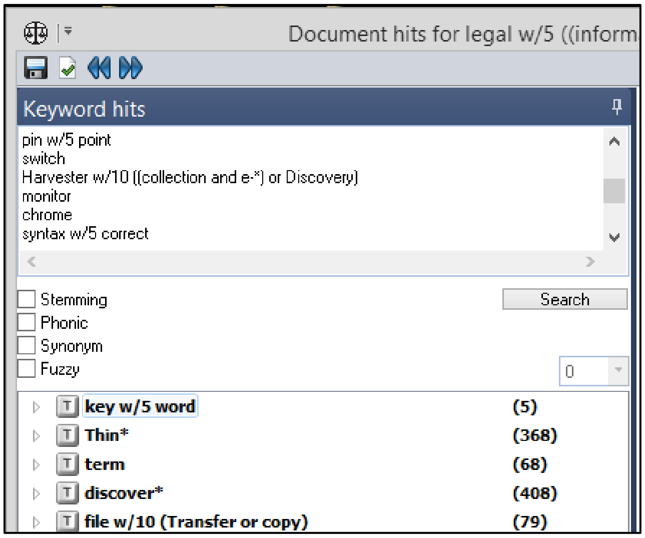
key w/5 word
□ DOJ searchmanual.pdf
□ Searching and Seizing Computers a
□ LegalJournalNovember2005.pdf
□ Bartek_Palak-Tutorial_handout_001
□ Bartek_Palak-Tutorial_handout.doc

the warrant.* In some cases, a "key word" search or similar surgical approach
may be possible.* Such an approach
may permit law enforcement to locate the incriminating files without conducting an
extensive search through
innocent files that happen to be mixed together with the incriminating files that are
the target of the search.* Notably,
the Fourth Amendment does not generally require such an approach.* See United
States v. Hunter, 13 F. Supp.2d
574, 584 (D. Vt. 1998) ("Computer records searches are no less constitutional
than searches of physical records,
where innocuous documents may be scanned to ascertain their relevancy."); United
States v. Lloyd, 1998 WL
846822, at *3 (E.D.N.Y. 1998).* However, in extensive dicta, the Tenth Circuit
has indicated that it favors such a
narrow approach because it minimizes the possibility that the government will be
able to use a narrow warrant to
justify a broader search. See United States v. Carey, 172 F.3d 1268, 1275-76,
1275 n.8. (10th Cir. 1999) (citing
Raphael Winick, Searches and Seizures of Computers and Computer Data, 8
Harv. J. L. &. Tech. 75, 108 (1994));
Campos, 221 F.3d at 1148.* See also Gawrysiak, 972 F. Supp. at 866
(suggesting in dicta that agents executing a
search for computer files "could have at the least checked the date on which each
file was created, and avoided
copying those files that were created before the time period covered by the

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.

<u>KEYWORD PREVIEW FEATURES</u>

The keyword preview window allows users to save keyword hit report by file list or category to an HTML or CSV file.



pin w/5 point
switch
Harvester w/10 ((collection and e-*) or Discovery)
monitor
chrome
syntax w/5 correct

key w/5 word	(5)
Thin*	(368)
term	(68)
discover*	(408)
file w/10 (Transfer or copy)	(79)

To access these options click on the disk icon in the upper left hand corner of the screen. After selection users will be prompted to browse to a location to store the file and provide a filename.

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.

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. This flag also forces the program to exit once the job is 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 then program after a job has been run. It is used in conjunction with the -as flag.

-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.

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 tasks. 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 the General tab**
- **Enter the commands in the text boxes provided under Scripting**

The command fields also support the following variables:

[SCDrive] – The drive letter that Harvester is running from (E:)
[JobName] – The name of the running job
[CName] – The name of the computer running the job
[UName] – The username of the logged-in user running the job
[Date] – The date the job was run (20-Mar-15)
[DateTime] – The date and time the job was run down to the second (20Mar15-110633)
[Logs] – The full path to the logs folder
[Target] – The full path to the target folder.

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 tasks. 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 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 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

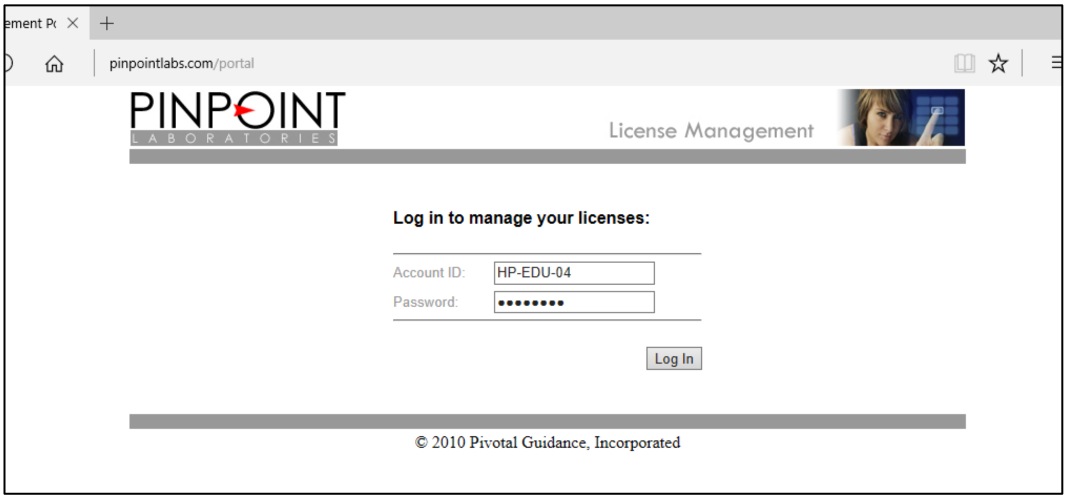
	emails that were responsive.
longpaths_source.log	This is a text file that contains any source paths that are greater than 255 characters.
longpaths_dest.log	This is a text file that contains any destination paths that are greater than 255 characters. These are logged because these files may be difficult 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
EnumExit.txt (PPLM)	Contains the exit conditions at the end of enumeration for a particular email store
CollectExit.txt (PPLM)	Contains the exit conditions at the end of processing for a particular email store
ErrorENUM.txt (PPLM)	Contains any errors encountered during the enumeration phase for a particular email store
ErrorCOLL.txt (PPLM)	Contains any errors encountered during the collection phase 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
_job.sdb (PPLCF)	Contains data for items encountered in a particular cloud source
_jobfile.scj (PPLCF)	Contains the settings and instructions for a particular cloud source
calling_command.txt (PPLCF)	Contains the instructions used to launch a particular cloud source
EnumExit.txt (PPLCF)	Contains the exit conditions at the end of enumeration for a particular cloud source
CollectExit.txt (PPLCF)	Contains the exit conditions at the end of processing for a particular cloud source
Error.txt (PPLCF)	Contains error information not related to debug functions within the processing of a particular cloud or email source
Log.txt (PPLCF)	Contains setup, version, and logging information for the enumeration and collection of a particular cloud source
MailComplete.txt (PPLCF)	Contains exit conditions for the last thread process for a particular cloud or email source
Progress.txt (PPLCF)	Contains the last progress message for the thread processing of a particular cloud source

Portable License Manager

There are numerous ways for the end-users to use portable license manager to their advantage.

Go to *pinpointlabs.com\portal* and log in with the Account ID and password included in your PLM Registration email from Pinpoint Labs.

Reference: Portable License Manager Training Video



Log in to manage your licenses:

Account ID: HP-EDU-04

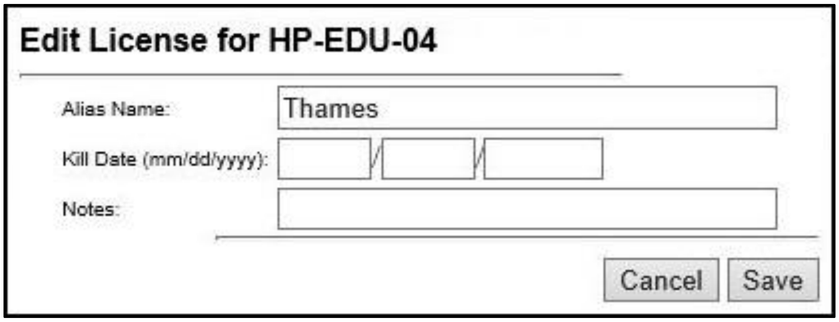
Password: ••••••••

© 2010 Pivotal Guidance, Incorporated

As an example, we have created a demo account with three licenses, and will use a combination of aliases and time-outs.

Click on the Edit License icon to create an alias. *An alias is used in place of the account ID* and is chosen by the user.

As an example I will use word 'Thames' as the alias.



Edit License for HP-EDU-04

Alias Name: Thames

Kill Date (mm/dd/yyyy): / /

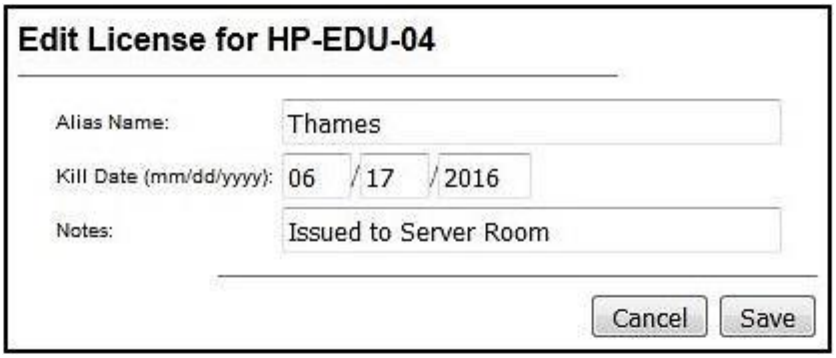
Notes:

Adding a time-out (or *kill date*) protects licenses from becoming inaccessible in the event of a system or drive failure, machine or drive re-purposing, or lost media.

The kill date can be as little as one day, and for effective license management we recommend no more than 15 days for portable drives.

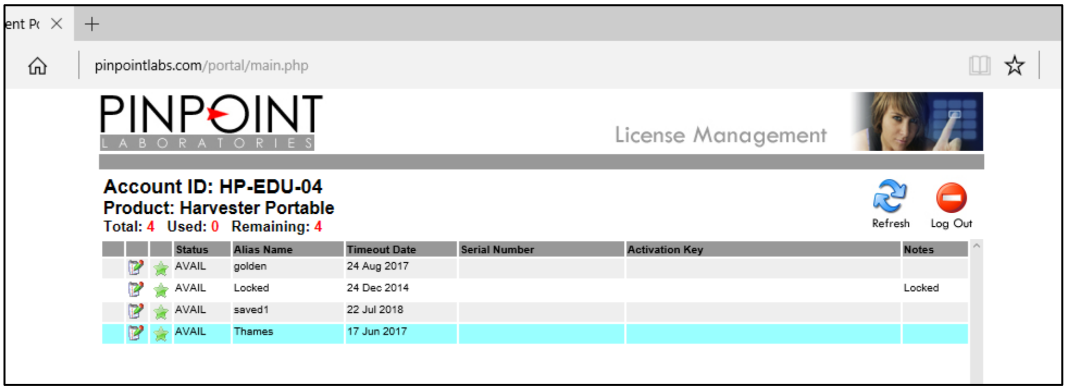
The same principal applies to activating directly on a computer. The Alias can be the name of the computer or department, and the kill date could be used as a safeguard the event of accidental deletion of the Harvester folder or a system crash on the PC.

Instructions or reminders can be placed in the Notes section. (This field is optional for license administrator and the end-user).



Edit License for HP-EDU-04

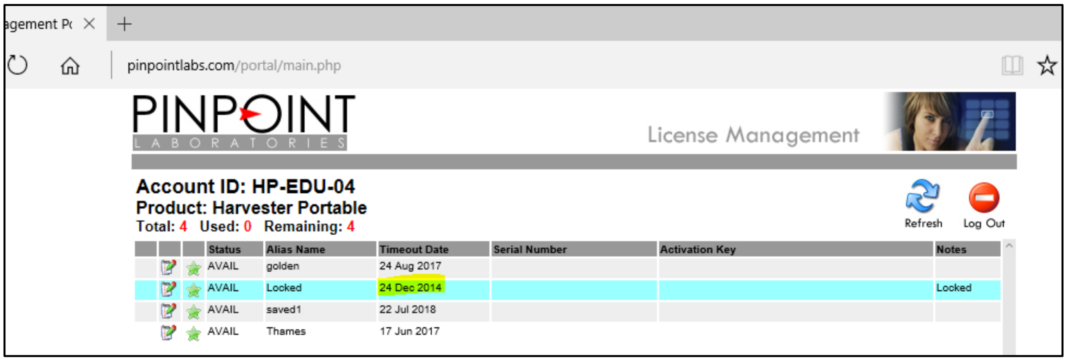
Alias Name:	Thames
Kill Date (mm/dd/yyyy):	06 / 17 / 2016
Notes:	Issued to Server Room



Account ID: HP-EDU-04
Product: Harvester Portable
Total: 4 Used: 0 Remaining: 4

			Status	Alias Name	Timeout Date	Serial Number	Activation Key	Notes
	📝	★	AVAIL	golden	24 Aug 2017			
	📝	★	AVAIL	Locked	24 Dec 2014			Locked
	📝	★	AVAIL	saved1	22 Jul 2018			
	📝	★	AVAIL	Thames	17 Jun 2017			

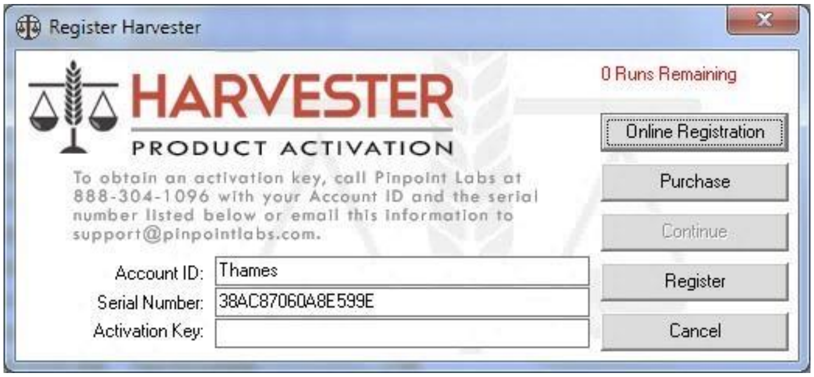
To prevent unauthorized use of an available license, entering an earlier date than present can keep the license from being activated until the date is changed.



Account ID: HP-EDU-04
Product: Harvester Portable
Total: 4 Used: 0 Remaining: 4

Status	Alias Name	Timeout Date	Serial Number	Activation Key	Notes
AVAIL	golden	24 Aug 2017			
AVAIL	Locked	24 Dec 2014			Locked
AVAIL	saved1	22 Jul 2018			
AVAIL	Thames	17 Jun 2017			

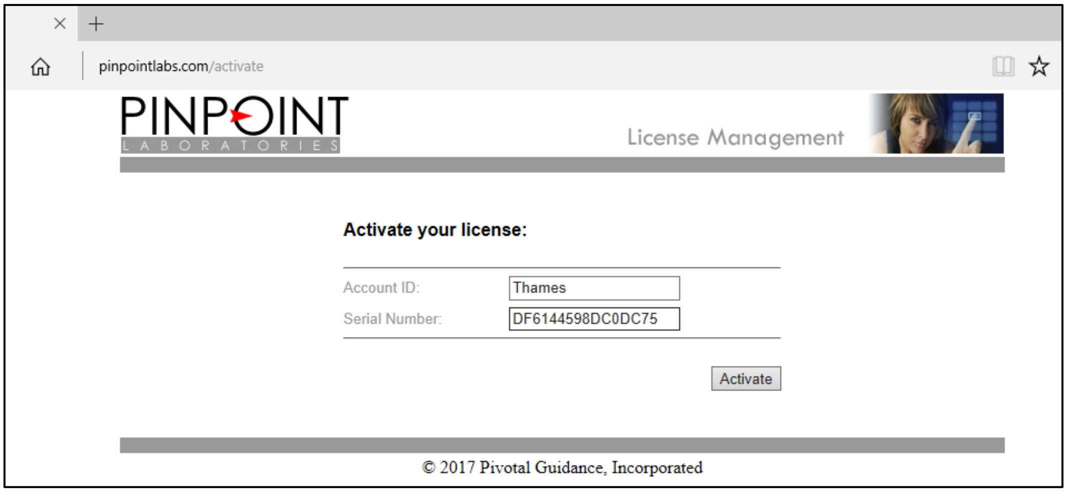
To activate, the user may use online activation, but if firewalls or lack of internet prevents online activation, the user has other options.



To obtain an activation key, call Pinpoint Labs at 888-304-1096 with your Account ID and the serial number listed below or email this information to support@pinpointlabs.com.

Account ID: Thames
Serial Number: 38AC87060A8E599E
Activation Key:

Log into pinpointlabs.com/activate on another computer or mobile phone, ***enter your alias*** from above and the serial number generated by SafeCopy or Harvester, and click Activate.

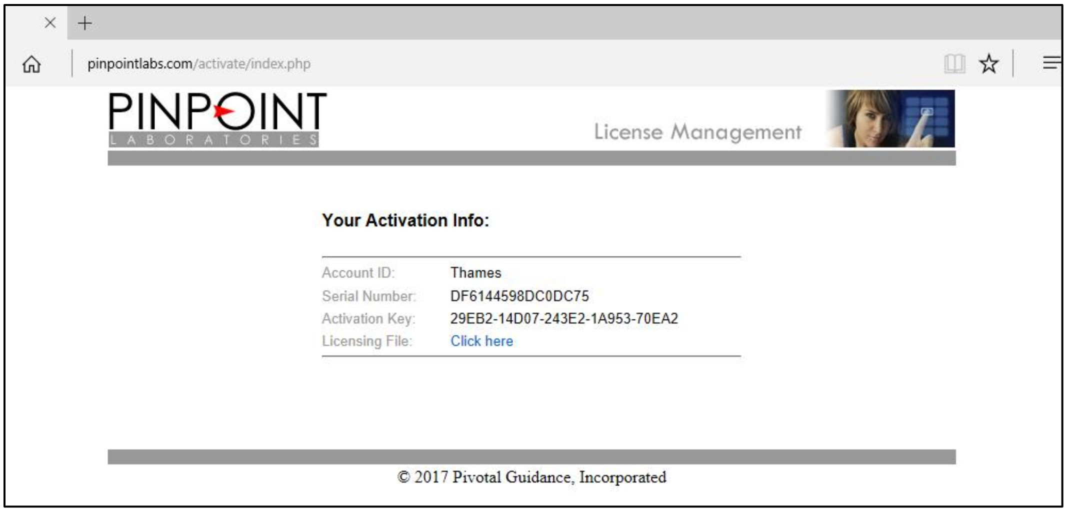


Activate your license:

Account ID: Thames
Serial Number: DF6144598DC0DC75

© 2017 Pivotal Guidance, Incorporated

Type or paste the activation code into the field provided, and click Activate to complete registration.



Your Activation Info:

Account ID:	Thames
Serial Number:	DF6144598DC0DC75
Activation Key:	29EB2-14D07-243E2-1A953-70EA2
Licensing File:	Click here

© 2017 Pivotal Guidance, Incorporated

NOTE: ***The PLM cannot deactivate an activated license***. Once a license is activated without a kill date, it must be manually deactivated.