Overview of the Central Repository

The Central Repository allows a user to find matching artifacts both across cases and across data sources in the same case. It is a combination of an ingest module that extracts, stores, and compares properties against lists of known or known bad properties, a database that stores these properties, and an additional panel in Autopsy to display other instances of each property. The Central Repository database can either be SQLite or PostgreSQL.

The following are some use cases for the Central Repository:

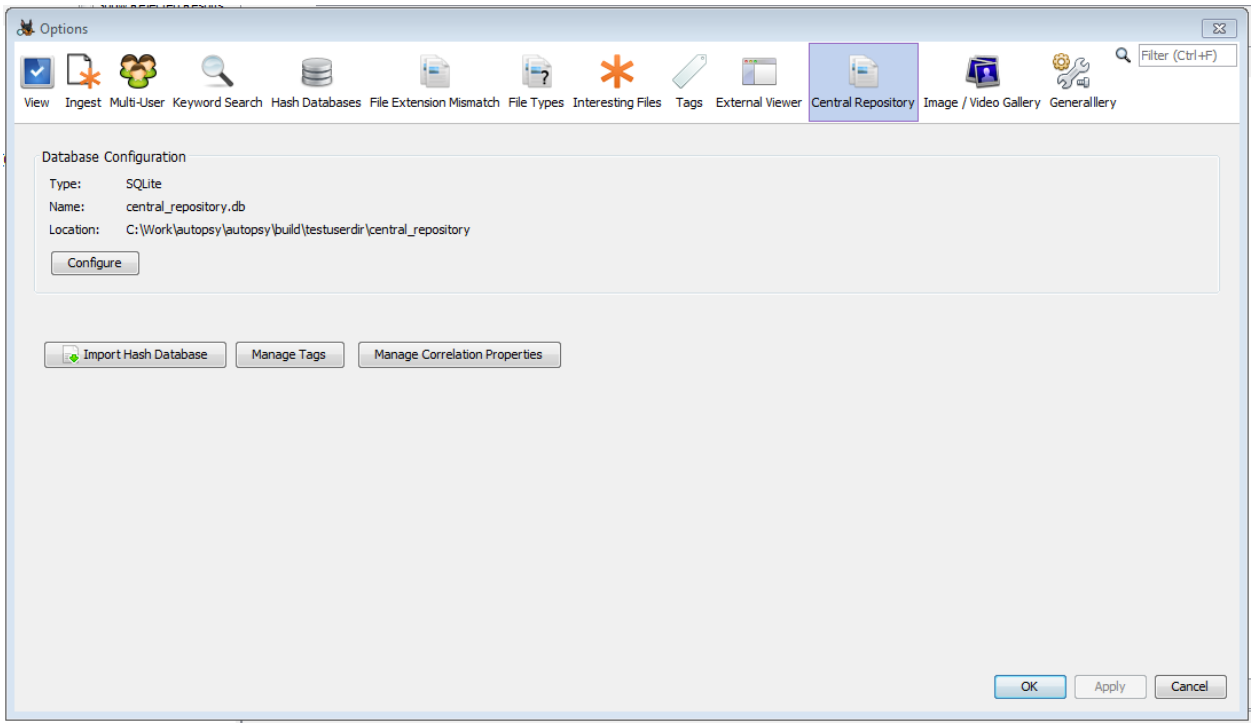
- Finding Other Instances of a Property
    - If you find a file or Autopsy artifact (such as a Web History item), there is a content viewer in the bottom right that will show you other cases that had this same file or that had items with the same feature (such as Domain name). You will also be able to see what other data sources in the same case had this feature.
- Alerting When Previously Notable Properties Occur
    - You can use the Central Repository to record which properties were associated with files and artifacts that were evidence (or notable). Once these properties have been tagged as “BAD” they will be added to the Interesting Items section of the tree when seen again in any future cases.
- Enabling a Global Hash Database
    - You can import hash databases into the Central Repository so that all Autopsy clients can use it instead of having local copies of the databases for each Autopsy client. You can do this for both "KNOWN" databases (i.e. NIST NSRL) and "KNOWN BAD"/notable databases.

Terms and Concepts

- **Central Repository** - The Autopsy feature containing the Central Repository Database and Correlation Engine Ingest Module. Also responsible for displaying correlated properties to the user
- **Central Repository Database** - the SQLite or PostgreSQL database that holds all the data
- **Correlation Engine Ingest Module** - The ingest module responsible for adding new properties to the database and comparing these properties against the Known/Known Bad lists
- **Property** - The data being stored/correlated. These can be file paths/MD5 hashes, email addresses, phone numbers, etc.

Setup

To start, open the main options panel and select the Central Repository icon.



If this icon is missing, perform the following steps:

- Select Tools->Plugins
- On the Installed tab, check the box next to CentralRepository then select the Activate button and go through the next few screens to activate the module

Setting up the Database

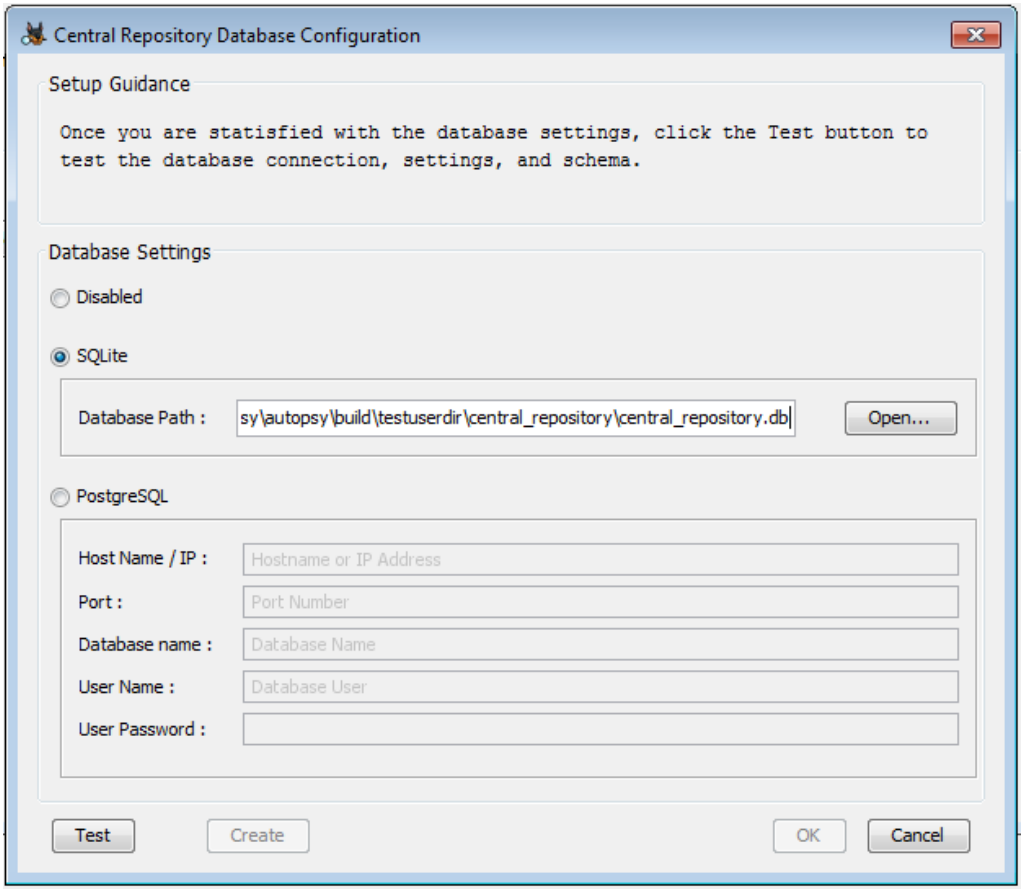
On the Central Repository options panel, click the Configure button to set up a database. There are three options here:

- Disabled
- SQLite - This option stores the database in a file. It should only be used when a single client will be accessing the database.
- PostgreSQL - This option uses a database server running either on the user's host or a remote server. This option must be used if multiple users will be using the same database.

Once a database has been configured, the three buttons on the main panel will be enabled, which will be described below.

Setting Up SQLite Deployment

There is only one step here, to specify the path and filename for the database. You can accept the default value or use the Open button to choose another path. The database file name can be called anything you want, but it is convenient to give it a ".db" suffix.



Once you have selected the path, click the Test button. If this is a new database or if the database is not found, you will see a red check and be prompted to click the Create button. If you see a green check next to the Test button, everything is ready to go. After clicking the create button, if you see a red check, there is a problem with the path you selected and you'll have to resolve that problem. Otherwise a green check will appear and you'll be able to save the settings.

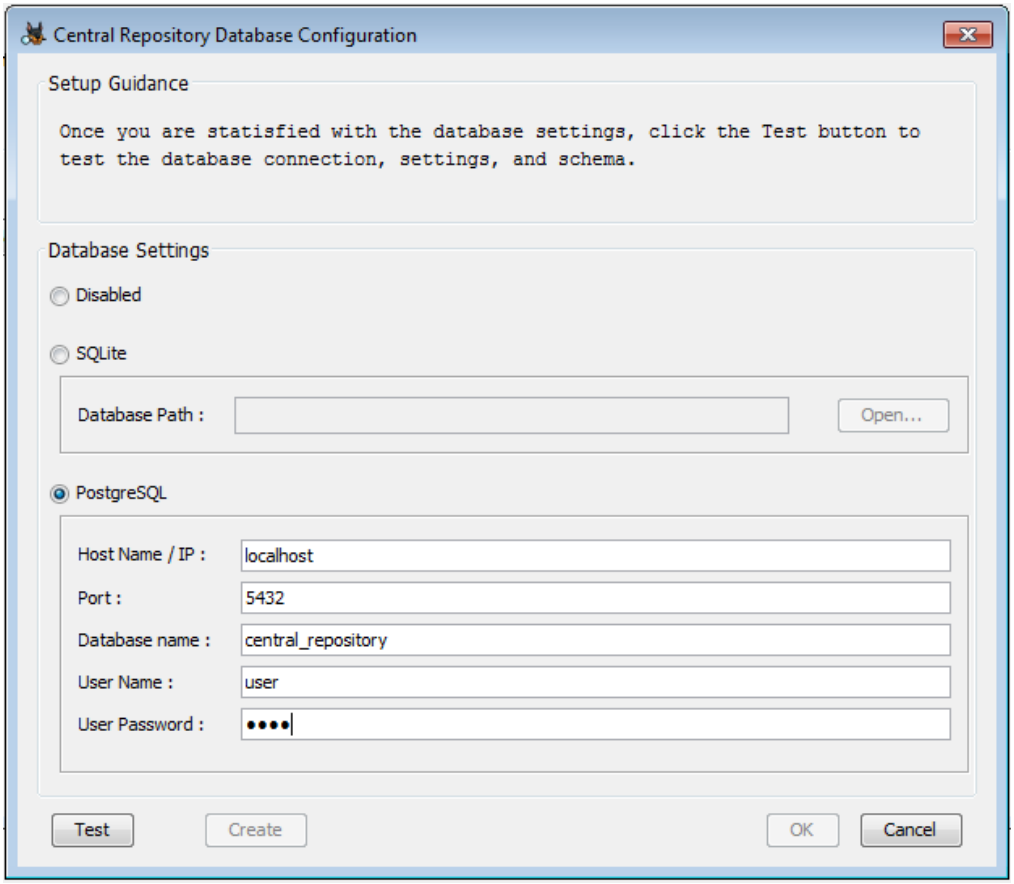
Once the test passes, click the OK button to save your selection and close the window.

Setting up PostgreSQL Deployment

If needed, see the Autopsy multi-user settings for help setting up your PostgreSQL server.

For PostgreSQL all values are required, but some defaults are provided for convenience.

1. Host Name/IP is the hostname or IP of your PostgreSQL server.
2. Port is the port that the PostgreSQL server is listening on; default is 5432.
3. Database name is the name of the database you are using for this module; default is central_repository.
4. User Name is the PostgreSQL user that owns and has full permissions to the database specified in step 3.
5. User Password is the password for the user.



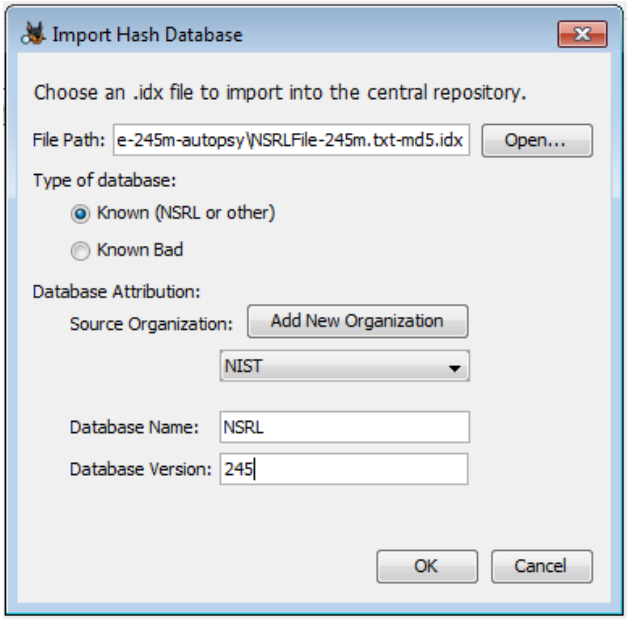
Once you have selected the path, click the Test button. If you see a green check next to the Test button, everything is ready to go. If the settings are incorrect, you will see a red check and be prompted to fix them. If this is a new database, you will see a red check and be prompted to click the Create button. After clicking the create button, if you see a red check, there is a problem with the database name or settings you selected and you'll have to resolve that problem. Otherwise a green check will appear and you'll be able to save the settings.

Once the test passes, click the OK button to save your selection and close the window.

Import Hash Database

The purpose of this feature is to store any Globally Known or Known Bad Artifacts in the database. Think of this feature like a dynamic Hash List. These artifacts are used during Ingest to flag files as Interesting. They are also displayed in the Content Viewer when a file or artifact is selected that is associated with one of the globally known artifacts.

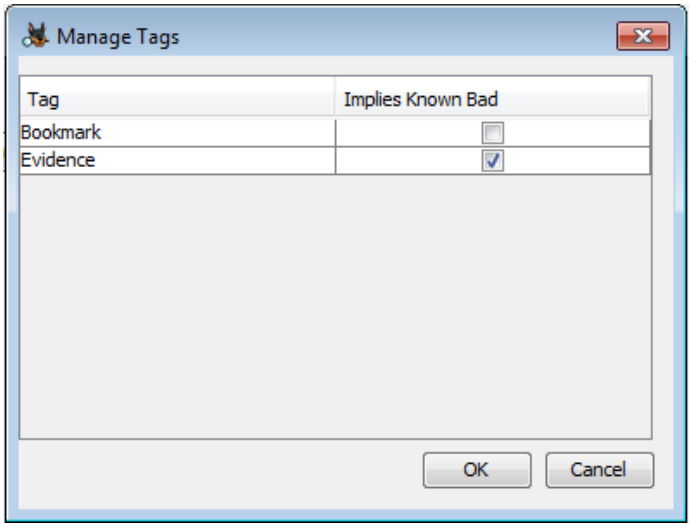
When importing a hash database, all fields are required. Current only .idx files are supported.



1. Select the Database Path using the Open button. This is the file containing the hash values that you want to import. You can import multiple files, but only one at a time. The format of these files must be the same format as used by the hash database module.
2. Select the database type. The type of content in the database being imported.
3. Define the attribution for this database.
    a. Select the Source Organization in the dropdown list. This is the organization that provided the hash database to you.
    b. If you do not see the Organization in the list, use the Add New Organization button to add it. Once you add it, you can then select it in the dropdown list.
    c. Enter a name for the dataset. This can be anything you want, but is often something like "child exploitation", "drugs", "malware", "corp hashlist", etc.
    d. Enter a version number for that dataset. This can be anything you want, but is often something like "1.0", "1.1a", 20170505", etc.
4. Click the OK button to start the import.

Manage Tags

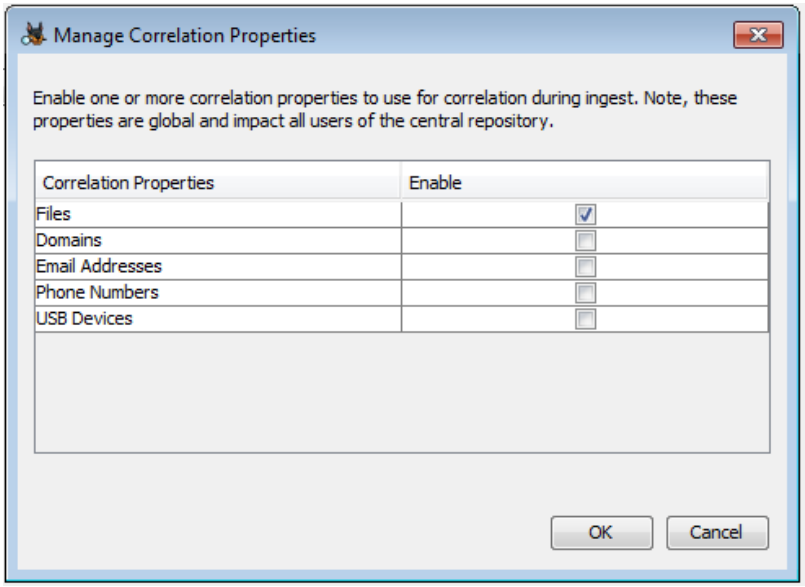
In Autopsy, you are allowed to define your own Tag names, tag files and artifacts,and add comments when you tag a file or artifact. The purpose of this feature is to associate one or more of those tags with this module to be used for Correlation. Associating a tag with the Correlation Engine means that when a file/artifact is tagged by the user, any property created from that file/artifact in the Correlation Database is marked as Bad. After this point, whenever the Correlation Engine Ingest Module creates a property that matches this Bad one, it is automatically flagged and added to the list of Interesting Items.



By default there is a tag called "Evidence" as the only tag associated with this module. To associate one or more tag(s) with this module, check the Correlate box next to the tag name(s) and click OK.

Manage Correlation Properties

The Correlation Engine ingest module can save different types of properties to the database. By default, only files are recorded, but this setting can be changed on the options panel through the Manage Correlation Properties button. Note that these settings are saved to the database, so in a multi-user setting any changes will affect all users.



Descriptions of the property types:

- Files
    - Files are correlated based on MD5 hash and file path and name. The Hash Database ingest module must be enabled.
- Domains
    - Domains are extracted from the various web artifacts, which primarily come from the Recent Activity module
- Email Addresses
    - Email addresses are pulled from Email Address hits from the Keyword Search module.
- Phone Numbers
    - Phone numbers are currently only extracted from call logs, contact lists and message, which come from the Android Analyzer module.
- USB Devices
    - USB device properties come from the registry parsing in the Recent Activity Module.

Using the Central Repository

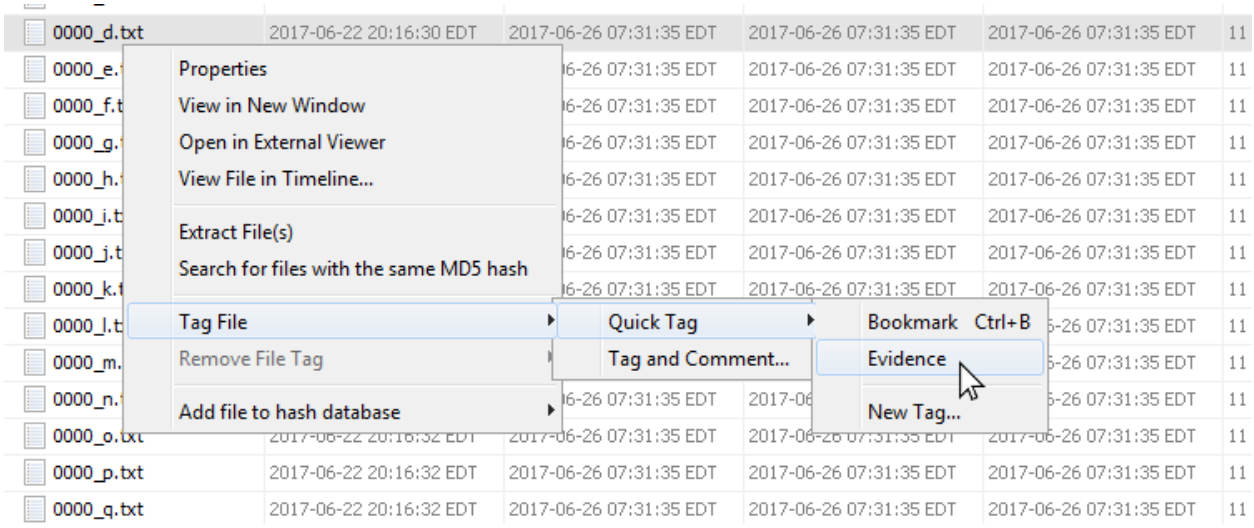
Ingest Module

The Correlation Engine ingest module is responsible for adding properties to the database and comparing each property against the list of Known/Known Bad property. It is best to run all ingest modules to get the most out of the Correlation Engine. For example, if Hash Lookup is not run then the Correlation Engine module will not put any files into the database. If the Correlation Engine module is not run on a particular case but the Central Repository is enabled,

there will still be some limited functionality. The Content Viewer will still display matching properties from other cases/data sources where the Correlation Engine was run.

Tagging Files

Any file or artifact that a user tags with one of the tags associated with the Correlation Engine will be added to the database as a file or artifact of interest. Any future data source ingest, where this module is enabled, will use those files or artifacts as if they were part of the Known Bad list, causing matching files from that ingest to be added to the Interesting Artifacts list in that currently open case.



Edit Central Repository Case Properties

By default, Autopsy lets you edit Case Details in the Case menu. When this module is enabled, there is an additional option in the Case menu, called "Central Repository Case Properties".

This is where the examiner can store a number of details about the case.

- The organization of the case examiner.
- The contact information of the case examiner.
- The case examiner's case notes.

To define the organization of the case examiner, simply select the organization name from the dropdown box. If the organization is not listed, you can click Add New Organization button. Once the new organization is added, it should be available in the dropdown box.

Viewing Results

Results from enabling the Central Repository and running the Correlation Engine Ingest Module can be seen in three places:

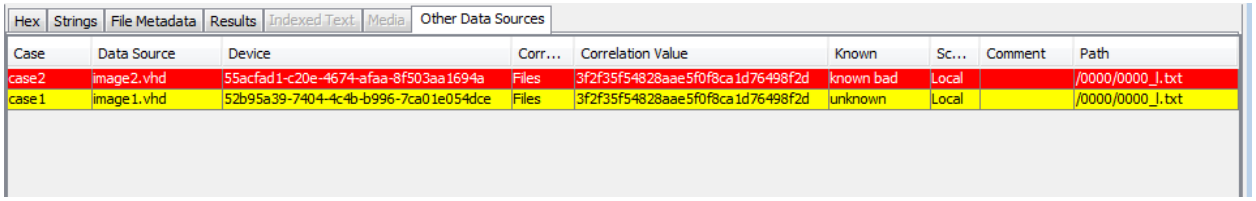
- The Content Viewer for each file or artifact will display all matching properties from other cases/data sources
- The Interesting Files node of the result tree will contain any files or results that matched properties previously flagged as BAD
- The Hashset Hits node of the result tree will contain any files that matched an imported Known Bad database entry

Content Viewer

The Content Viewer panel is where previous instances of properties are displayed. This module adds a new tab to the Content Viewer. The tab for this module is called "Other Data Sources". It can display data that is found in other cases, other data sources for the same case, or imported global artifacts.

If at least one other case or data source has been ingested with this module enabled, there is a potential that data will be displayed in the Other Data Sources content viewer. Note that the Correlation Engine Ingest Module does not have to have been run on the current data source to see correlated files from other cases/data sources. If the selected file or artifact is associated by one of the supported Correlation Types, to one or more file(s) or artifact(s) in the database, the associated files/artifacts will be displayed. Note: the Content Viewer will display ALL associated files and artifacts available in the database. It ignores the user's enabled/disabled Correlation Properties.

By default, the rows in the content viewer will have background colors to indicate if they are known to be of interest. Files/artifacts that are Known Bad will have a Red background, Unknown will have Yellow background, and Known will have a White background.



The user can click on any column heading to sort by the values in that column.

If the user right-clicks on a row, a menu will be displayed.
This menu has several options.

1. Select All
2. Export Selected Rows to CSV
3. Show Case Details
4. Show Commonality Details

Select All

This option will select all rows in the Content Viewer table.

Export Selected Rows to CSV

This option will save ALL SELECTED rows in the Content Viewer table to a CSV file.
By default, the CSV file is saved into the Export directory inside the currently open Autopsy case,
but the user is free to select a different location.

Note: if you want to copy/paste rows, it is usually possible to use CTRL+C to copy the selected rows and then CTRL+V to paste them into a file, but it will not be CSV formatted.

Show Case Details

This option will open a dialog that displays all of the relevant details for the selected case. The details will include:

- Case UUID
- Case Name
- Case Creation Date
- Case Examiner contact information
- Case Examiner's notes

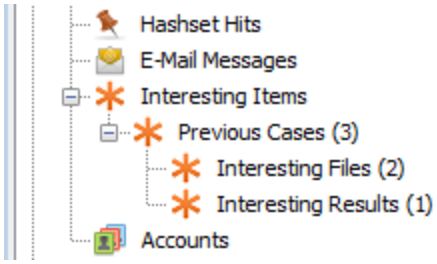
These details would have been entered by the examiner of the selected case, by visiting the Case -> Central Repository Case Properties menu, when that case was open.

Show Commonality Details

The concept of Commonality simply means, how common is the selected file. The value is the percentage of case/data source tuples that have the selected file or artifact.

Interesting Items

In the Results tree of an open case is an entry called Interesting Items. When this module is enabled, all of the enabled Correlatable Properties will cause matching files to be added to this Interesting Items tree during ingest.



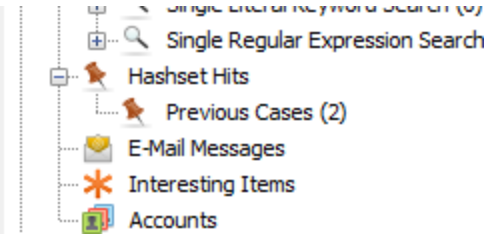
As an example, if the Files Correlatable Property is enabled, and the ingest is currently processing a file, for example "badfile.exe", and the MD5 hash for that file already exists in the database as a KNOWN BAD file, then an entry in the Interesting Items tree will be added for the current instance of "badfile.exe" in the data source currently being ingested.

The same type of thing will happen for each enabled Correlatable Property.

In the case of the phone number correlatable type, the Interesting Items tree will start a sub-tree for each phone number. The sub-tree will then contain each instance of that Known Bad phone number.

Hashset Hits

Matches from any imported hash databases will be displayed in the Hashset Hits section of the results tree.



In the Other Data Sources tab, imported hash set matches are marked with “Global” scope and do not have case/data source information.