Revision History

The following changes have been made to this document.

<table>
<thead>
<tr>
<th>VERSION NUMBER</th>
<th>REVISION DATE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2</td>
<td>October 2015</td>
<td>Minor Improvements</td>
</tr>
<tr>
<td>7.2</td>
<td>December 2016</td>
<td>Minor changes</td>
</tr>
</tbody>
</table>

DISCLAIMER

© 2016 Nuix. All rights reserved.

This publication is intended for informational purposes only. The information contained herein is provided “as-is” and is subject to change without notice. Although reasonable care has been taken to ensure that the facts stated in this publication are accurate and that the opinions expressed are fair and reasonable, no representation or warranty, express or implied, is made as to the fairness, accuracy or completeness of the information or opinions contained herein, and no reliance should be placed on such information or opinions. Neither Nuix nor any of its respective members, directors, officers or employees nor any other person accepts any liability whatsoever for any loss arising from any use of such information or opinions or otherwise arising in connection with this publication. Furthermore, this publication contains the confidential and/or proprietary information of Nuix which may not be reproduced, redistributed, or published in any form or by any means, in whole or in part, without the express prior written consent of Nuix. The use, reproduction, and/or distribution of any Nuix software described in this publication requires an applicable software license.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workbench Tab</td>
<td>4</td>
</tr>
<tr>
<td>Statistics Tab</td>
<td>7</td>
</tr>
<tr>
<td>Word List Tab</td>
<td>8</td>
</tr>
<tr>
<td>Addresses Tab</td>
<td>9</td>
</tr>
<tr>
<td>History Tab</td>
<td>10</td>
</tr>
<tr>
<td>Global Options</td>
<td>11</td>
</tr>
<tr>
<td>Search</td>
<td>11</td>
</tr>
<tr>
<td>Document Navigator</td>
<td>12</td>
</tr>
<tr>
<td>Results</td>
<td>12</td>
</tr>
<tr>
<td>Context</td>
<td>12</td>
</tr>
<tr>
<td>Preview</td>
<td>12</td>
</tr>
<tr>
<td>Tagging</td>
<td>13</td>
</tr>
<tr>
<td>Default Tabs</td>
<td>13</td>
</tr>
<tr>
<td>Metadata Profiles</td>
<td>13</td>
</tr>
<tr>
<td>Custom Metadata Templates</td>
<td>14</td>
</tr>
<tr>
<td>Digest Lists</td>
<td>14</td>
</tr>
<tr>
<td>Shingle Lists</td>
<td>14</td>
</tr>
<tr>
<td>Word Lists</td>
<td>14</td>
</tr>
<tr>
<td>Fuzzy Hash Lists</td>
<td>15</td>
</tr>
<tr>
<td>Fuzzy hash lists &amp; filtering</td>
<td>16</td>
</tr>
<tr>
<td>Configuration Profiles</td>
<td>16</td>
</tr>
<tr>
<td>Saved Searches</td>
<td>16</td>
</tr>
<tr>
<td>Search Macros</td>
<td>16</td>
</tr>
<tr>
<td>Getting started with .macro file</td>
<td>16</td>
</tr>
<tr>
<td>Run Search Macro in Nuix</td>
<td>17</td>
</tr>
<tr>
<td>Launching</td>
<td>17</td>
</tr>
<tr>
<td>Imaging Profiles</td>
<td>17</td>
</tr>
<tr>
<td>Production Profiles</td>
<td>17</td>
</tr>
<tr>
<td>Memory</td>
<td>17</td>
</tr>
<tr>
<td>Hot Folder</td>
<td>17</td>
</tr>
<tr>
<td>Searching Within Evidence</td>
<td>18</td>
</tr>
<tr>
<td>Search Bar</td>
<td>18</td>
</tr>
<tr>
<td>Per-Day Timeline View</td>
<td>23</td>
</tr>
<tr>
<td>Timeline View</td>
<td>23</td>
</tr>
<tr>
<td>Entities View</td>
<td>23</td>
</tr>
</tbody>
</table>
Files View........................................................................................................... 24
Addresses View.................................................................................................. 24
Shingles View .................................................................................................... 24
Words View ........................................................................................................ 24
Topics View ........................................................................................................ 24
Event Map View ................................................................................................ 24
Network View .................................................................................................... 25
Map View ............................................................................................................ 26
Hide Immature Items Option ............................................................................ 26
Deduplicate Results Option ............................................................................. 27
Preview Toolbar ................................................................................................. 28
Preview Item Context ........................................................................................ 28
Preview Item Detail Tabs ................................................................................... 29
Advanced Query Builder .................................................................................. 32
Appendix A: Nuix Metadata Fields ................................................................... 38

Figures

Figure 1: Workbench Tab .................................................................................... 5
Figure 2: Statistics Tab ......................................................................................... 7
Figure 3: Word List Tab ....................................................................................... 8
Figure 4: Addresses Tab ..................................................................................... 9
Figure 5: History Tab .......................................................................................... 10
Figure 6: Global Options ................................................................................... 11
Figure 7: Resemblance Threshold .................................................................... 12
Figure 8: Default Tabs ....................................................................................... 13
Figure 9: Metadata Profiles ............................................................................... 14
Figure 10: Nuix Defined Metadata .................................................................... 15
Figure 11: Fuzzy Hash List Filtering ................................................................. 16
Figure 12: Preview Pane ..................................................................................... 28
Figure 13: Preview Item Details ........................................................................ 29
Figure 14: JSON File Format ............................................................................ 36
Workbench Tab

The first time you use Proof Finder and have processed data, you are presented with the Workbench tab. This tab hosts the primary tasks of excluding, filtering, and searching for data within the case. You can also analyze data, preview individual items, and tag items from this tab.

To create a new case and load data or open an existing case, refer to our Proof Finder Reference Guide - Working with Cases and Loading Data.

Figure 1: Workbench Tab

1. Evidence Navigator - located within the Document Navigator pane on the Workbench tab. It displays the data loaded into the case in its original source folder hierarchy, allowing you to sort evidence based on item name, browse the evidence by folder, or filter the set of files to view or analyze.

2. Excluded Items Navigator - located within the Document Navigator pane on the Workbench tab. It displays the items that you have chosen to exclude from all future actions within the case, such as searching, reviewing, analyzing, and tagging. Excluded items will not appear in the result sets you generate.

3. Custodians Navigator - located within the Document Navigator pane on the Workbench tab. It displays all the custodians created within the case and the items that have been assigned to them.

4. Item Sets Navigator - located within the Document Navigator pane on the Workbench. It allows the management of deduplication across sets of documents within the same case. Please note that Item set navigator is hidden until an item set is created.

5. Filtered Items Navigator - located within the Document Navigator pane on the Workbench tab. It displays built-in Nuix filters, which are based on metadata gathered from ingested items. You can view a breakdown of the items in the collection based on a variety of metadata criteria, such as item type, email attachments, irregular items, annotated items, skin toned images, tagged items, review jobs, and more.
6 **Document Navigator filter** - located at the base of the Document Navigator pane on the Workbench tab. This filter aids to show/hide selected navigators for ease of interrogating data sets, by allowing users to customize the view and use it repeatedly across subsequently opened workbenches.

7 **Search Bar** - performs simple and complex searches against the evidence set. Click **Advanced** search option on the search bar to select the types of metadata you need, add keywords, values, or other criteria. The advanced tool builds the search syntax for you, allowing you to edit/remove parts of the expression as you work. To clear a search query, click **Clear**. You can also specify a date range by selecting the **Date Filter** drop down menu and enter the date range.

8 **Results Pane** - displays a list of the items that match your selected criteria, whether from filters in the Document Navigator, a search, or other operations. By default, the view of the items is in a tabular list format, showing the metadata in columns from the associated metadata profile applied to that view. At the bottom of the selected view, Proof Finder tells you how many items are displayed that match the criteria. If more items exist than can be shown in the view, it will state that. It also shows how many items are selected, if any and how many items are removed from the list because they are immaterial or duplicates, if you use those options. You can also alternatively choose to view the items by Statistics or by Visualizations from the **View By** menu.

9 **Preview Pane** - is comprised of information and tools that allow you to view the item individually, including the metadata associated with the item and additional information to help analyze the context of the item.

10 **Review and Tag Pane** - allows you to create and apply tags to the items you review in the Preview pane. You can assign numerical values for up to nine tags to support efficient tagging of large result sets from the keyboard. You can also apply tags in bulk to family members or duplicates, when reviewing and tagging an item. To associate a tag with that numerical value, right-click on the numerical value, select **Assign tag to this Shortcut**, and select the tag to be associated.

**Notes:**
- To set or change the tabs you would like to automatically open after loading a case or creating a case, navigate to **File > Global Options > Default Tabs** and select the desired tabs.
- Create a new Workbench tab by navigating to **Window > New Workbench Tab**.
- To add other tabs, navigate to **Reports > New (Statistics, Word, Addresses, History, Fast Review Statistics) tab**, a new tab is opened.
- To close a tab, right-click on the tab and select **Close, Close Others** or **Close All**.
- Proof Finder creates new Workbench tabs as required when certain actions are performed by drilling down to result sets (for example selecting similar items or duplicate links).
- The tabs provide flexible layout to aid in the analysis and review of data by allowing any pane such as the Preview Pane in the Workbench tab to be moved to a second monitor for side by side analysis.
Statistics Tab

The Statistics tab offers an overview of all file types processed in the case and their respective frequency within the dataset, including a list of raw file extensions found and files classified as irregular files. The Statistics tab should be carefully reviewed after you load data into a new case, subsequently each time you add evidence to a case, to gain a full understanding of your ingested data and any exceptions encountered which might need further analysis.

Open a new Statistics tab by navigating to Reports > New Statistics Tab.

![Statistics Tab](image)

Figure 2: Statistics Tab

The Statistics tab is divided into three main areas, Processed Files, Raw File Extensions, and Irregular Files. You can perform the following operations within the Statistics tab:

- Open a new workbench containing items for a specific file type by double-clicking on any row in the Statistics tab.
- Sort a column in ascending or descending order by single-clicking in the column header. The default order is set to ascending.
- Export the Statistics view by using File > Export > Export View.

Note: To open a statistics view on a selected set of data, select the required items to report on in the Results pane and select View By Statistics.
Word List Tab

The Word List tab provides a listing of every word that appears in the content and properties of the data set or a custom word list, and a count of the number of items containing that word. Open a new Word List tab by navigating to Reports > New Word List Tab.

You can use the following functions in the Word List tab:

- Use the drop-down list on the upper left to display words by a word list. The default setting is ASCII Words, which displays a list of all words that are ASCII based the data set. You can import a text file containing a custom word to scope the listing on this tab to only those words that are of interest.
- Use the drop-down menu to search across all fields, item content or properties of all items.
- Type one or more characters into the Filter text box to filter the list that displays to match your entry. This filter is based on an anchor at the beginning of the word, so "ranteed" will not show "guaranteed". The filter supports numbers, letters and symbols.
- Open a result set containing only the items that include a specific word by double-clicking the row that contains the word.

Notes:

- Proof Finder views a word as any item that is surrounded by white space so 24014 is considered a word. From a practical perspective this information can be trivial or it could be critical as a zip code.
- All words are listed, including all character sets and symbols.
- To open a word list view on a selected set of data, select the required items to report on in the Results pane and select View By Word List.
Addresses Tab

The Addresses tab provides a listing of every email address that appears in the properties of the data set, and a count of the number of items containing that address. Open a new Addresses tab by navigating to Reports > New Addresses Tab.

![Figure 4: Addresses Tab](image)

You can use the following functions in the Addresses tab:

- Show the results grouped by domain group or expanded to show all addresses.
- Filter the results by type of correspondence to only show emails From, To, CC, or BCC.
- Find matching results to a particular domain group or user by using the Find function from the Edit menu.
- In a new Workbench tab containing only the items that include a specific word by double-clicking the row that contains the word.

Notes:

- Proof Finder views an address as any item that is removed from the transport headers of email items. From a practical perspective this could be just a name from an address book or a fully resolved email address.
- To open an addresses view on a selected set of data, select the required items to report on in the Results pane and select View By Addresses.
History Tab

The History List tab provides a log of a variety of events and user actions in the case, providing details of when the case was opened, searches that were performed, and when and who annotated items. Timestamps are recorded for each event, along with who performed the event, the type of event, the status of the event, and event details. Open a new History tab by navigating to Reports > New History Tab.

![Figure 5: History Tab](image)

Actions you can perform from this tab include:

- Re-run a specific search query by double-clicking on a search event. A new workbench tab displays showing the results of the query against the current data set. This is not a static result set, so if new evidence is added to a case, the number of results will reflect the new evidence.
- Filter the History results by the type of event performed, by the user that performed the event or by the date the event was performed.
- Sort the columns in ascending or descending order by single-clicking on a column header. The default order is set to ascending.
- Export the contents of the History tab by using File > Export > Export View.
Global Options

Configure global options by navigating to File > Global Options.

**Search**

Search allows you to perform searches by selecting:

- **Search Content** - allows searches performed to find results in the Content or text only of the indexed documents.
- **Search Properties** - allows searches performed to find results in the Properties section only of the indexed documents.
- **Search Names** - allows searches performed to find results in the Names field only of the indexed documents.
- **Search Path Names** - allows searches performed to find results in the Path Names field only of the indexed documents.
- **Search Evidence Metadata** - allows searches performed to find results in the Evidence Metadata or user defined data section only of the indexed documents.

**Resemblance threshold**: Set the level of similarity required to allow documents to be found as near duplicates of each other. The resemblance value ranges between 0 and 1 with 1 representing very similar documents. The default value is set at 0.5.
Figure 7: Resemblance Threshold

**Document Navigator**

**Document Navigator** allows you to enable certain criteria including:

- Enable “filter excludes all” warning
- Enable evidence panel count updates
- Enable excluded items panel count updates
- Enable item set panel count updates
- Enable production set panel count updates
- Enable automatic classifier panel count updates
- Enable - allows you to enable a number of additional count update modes

**Results**

**Results** allows you to set results options and viewer limits.

**Results Options:**

- Show the ancestry and annotation column
- Show immaterial items in results by default

**Viewer Limits:**

- Results table row limit
- Networks node limit

**Context**

**Context** allows you to manage context graph options, limits, and manage custom analysis links.

**Preview**

**Preview** allows you to set preview options and highlighting options.

- Enable duplicates, similar and related counts
- Enable cluster counts
- Break phrases into words when highlighting
Tagging
Tagging allows you to select to enable family and duplicate counts.

Default Tabs
Default Tabs allows you to set the tab you wish to view by default in Proof Finder when opening a new case.

- **Workbench** - This tab hosts the primary tasks of excluding, filtering, and searching for data within the case. You can also analyze data, preview individual items, and tag from this tab. This tab is set to display by default when you open a case.
- **Statistics** - This tab displays information about the processed and irregular files by file type, including number processed, corrupted, and encrypted, as well as a percentage of each file type encountered.
- **Search and Tag** - This tab allows you to search processed data using queries, and tag the results for future reference.

![Figure 8: Default Tabs](image)

Metadata Profiles
Metadata Profiles allows you to manage the presentation and export of metadata. You can create metadata profiles for specific item types (email/files), specific purposes (exception handling), or the specific load file formats required. The Default Metadata Profile is provided with the application.
Custom Metadata Templates

Custom Metadata Templates allows you to create templates for custom metadata.

Digest Lists

Digest Lists allows you to import digest lists from third party sources as well as directly create them from within Proof Finder. The plain text format includes a single digest per line. When creating plain text hashes, ensure that there is no trailing punctuation or whitespace.

Digest lists are used to:
- Eliminate system files or other application files that have known signatures and little or no value to the investigation. This process is called “De-NISTing”.
- Eliminate or suppress inappropriate content. If inappropriate content is detected, you can import/generate a hash list of known inappropriate content, and pass that along as part of the export process to allow this content to suppress downstream.

Shingle Lists

Shingle Lists allows you to create a shingle list from a set of key documents that you can use as a filter against the dataset or import into other cases to use against other datasets. You can select one or more shingle lists in the Filtered Items pane to return a list of items that are similar within the resemblance threshold that you have set in the Search Options section.

Word Lists

Word Lists allows you to import a .txt file containing a list of keywords that you can use as a filter against the dataset. You can select one or more word lists in the Filtered Items pane to produce a list of items that include the words you have compiled.

Each new word in the text file must be placed on a separate line. There is no limit to the number of words that
you can include in the word list, but the greater the number of words in a list, the greater the number of matching documents you will receive in the Results list.

**Notes:**
- Multiple words on a single row are treated as an exact phrase. (e.g. Dog Cat Mouse, are treated like a search for “dog cat mouse”). Quotes are unnecessary, and will be stripped.
- Boolean or other searches are not supported within a word list, so “(classification OR maxim)” is not valid. To perform a series of Boolean or complex searches against a Proof Finder dataset, the scripting interface provides you with a means of automatically executing queries, and applying classifications to the result set. If complex queries or reporting is required, see the scripting section for additional detail.

**Fuzzy Hash Lists**

Fuzzy hashing is a method in which the SSdeep program is used to generate a hash value of a file. This hash value can then be used to search and identify similar files to the original hashed file, producing a percentage of similarity.

Typically, investigators process their case evidence generating a MD5 hash of each item. This can then be used for verification to ensure that the item has not been modified, but also to identify files that are exactly the same. When comparing the hash values of two files, the outcome is either, match or no match. This only indicates that the file is not an exact match, changing one byte of the file can cause this. By generating a fuzzy hash for items, investigators can identify items that have a similar binary content. This is useful while identifying malware and the family it may belong to when the code was modified so it will not be responsive to traditional anti-virus scan.

Fuzzy hash value can be set in the Evidence Processing settings by selecting the SSDeep option under **Data Processing Settings > Digest Settings**. Once the case has finished processing, the investigator can identify the fuzzy hash of an item from the Nuix-defined Metadata list under the **Preview Pane > Metadata** tab.

![Figure 10: Nuix Defined Metadata](image)

Alternatively, user can use the following fuzzy hash syntax on the search bar.

<table>
<thead>
<tr>
<th>Fuzzy Hash Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>fuzzy-hash:value</td>
<td>Enter a fuzzy hash value to identify an exact match</td>
</tr>
<tr>
<td>fuzzy-hash:value;score</td>
<td>Enter a fuzzy hash value and minimum threshold score to match against</td>
</tr>
<tr>
<td>fuzzy-hash-list:name</td>
<td>Enter the fuzzy hash list name to identify all files that match the hashes exactly</td>
</tr>
<tr>
<td>fuzzy-hash-list:name;score</td>
<td>Enter the fuzzy hash list name and minimum threshold score to match against exactly</td>
</tr>
</tbody>
</table>

For syntax where you wish to define a score this can also be a range for example `fuzzy-hash-list:name:[75-99]`. When entering fuzzy hashes into the search bar you must escape the colons (`\`)
Fuzzy hash lists & filtering
Fuzzy hash lists can be imported through the Global Options > Fuzzy Hash Lists or exported from a group of selected items by right-clicking in the Results pane > Export > Export Fuzzy Hash List. Once you have Fuzzy hash list either imported or exported into a list the user can then use filtering pane to identify similar items based on a Fuzzy hash scoring.

Expanding the Fuzzy Hash list filter updates the responsive count. Expanding each fuzzy hash list in case will then break down the overall count into fuzzy hash list and then into a scoring threshold of either High, Medium, or Low.

- **High** - Item is between 77 & 99 percent similar to an item within that fuzzy hash list.
- **Medium** - Item is between 40 & 69 percent similar to an item within that fuzzy hash list.
- **Low** - Item is between 1 & 39 percent similar to an item within that fuzzy hash list.

![Fuzzy Hash List Filtering](image)

**Figure 11: Fuzzy Hash List Filtering**

Configuration Profiles
Configuration Profiles allows you to select profiles to be used as runtime configurations.

Saved Searches
Saved Searches allows you to select and import saved searches.

Search Macros
Search Macros enables a repeatable process of identifying relevant items, screening them for sensitive information prior to exporting from Nuix. It allows investigators to build a profile around individuals being investigated. This is portable from case to case.

Getting started with .macro file
The following steps help you get started with .macro file.

- In a text editor, enter the syntax of your query, e.g., ((raptor OR enron) w/3 balance) AND "this transaction". The macro can contain any query you enter in the search bar, including regexes.
- Save the file to C:\Users\XXX\AppData\Roaming\Nuix\Search Macros directory with a .macro extension, e.g., Searchterm.macro.
  
  **Note:** Ensure you specify a simple name for the .macro file. If it contains Unicode text, you may need to use UTF-8 text encoding without a byte order mark (BOM).
Select from a list of available search macros to identify relevant items and screen them for routine information. Macros may be used if the screening process needs to be repeated regularly.

To import a search macro, navigate to **Global Options > Search Macros** and select **Import** to import a (.macro) file. Search macros can currently be imported in Proof Finder export format.

To make the .macro file available to other users, save the file to C:\ProgramData\Nuix\Search Macros.

**Run Search Macro in Nuix**

You can run search macro by entering ${macro} syntax in the search field. e.g., $searchterm. The results pane displays the search macros once you run the search. To view all available search macros, navigate to **Filtered Items > Search Macros**.

If you wish to edit the search, you must edit the .macro file. You will not be able to append the $searchterm in the search field to retrieve items that are not listed in the .macro file.

**Launching**

**Launching** allows you to set the default native application Nuix uses to open email messages by setting the email format.

**Imaging Profiles**

This option is available only for eDiscovery license users. It provides a means to set granular rules to the documents that are imaged by mime types when creating images, whether populating stores or exporting. It allows you to save the settings consistently by adding imaging profiles where you can set the rendering options, add additional slipsheet templates and set rules for imaging content.

**Production Profiles**

This option is available only for eDiscovery license users. It allows you to add production profiles by customizing the headers and footers style, setting the Text File export settings and export rules.

**Memory**

**Memory** allows you to configure the amount of RAM you wish to allocate to Proof Finder. The amount of RAM allocated to Proof Finder can be accustomed based on the circumstance and current use case. It is recommended to allocate at least 4GB to Proof Finder.

**Notes:**

- Before starting to analyze the case data, ensure the Proof Finder has been allocated enough memory.
- The maximum memory that can be allocated on a 32-bit OS is 1300 MB. If you are unable to set the value higher, confirm that you are using a 64-bit OS.

**Hot Folder**

**Hot Folder** allows you to automatically import PDFs into Nuix in a specified folder and allows for checking for hot folder updates.
Searching Within Evidence

After you have created a case and ingested case data, you can search through the evidence to find information that is of interest. Nuix offers the following ways to quickly search through the body of evidence:

Use the Search bar at the top of the window to search using simple keyword and Nuix search syntax queries and/or dates.

- Use the predefined Filtered Items categories to refine the results set based on preset Nuix search syntax for particular metadata.
- Use the Advanced Search tool to build and run more complex queries.

Search Bar

The Search bar, located at the top of the window, provides you with a tool for performing both simple and complex searches against the evidence set. For more information, refer to the Searching within Evidence section. Searches run against items that match any existing filters and items that are not excluded.

The Search bar comprises the following components.

- **Previous and Next** - advances backwards and forwards through the searches already performed in the currently open session of Nuix. Searches performed prior to the current session are not available. When you use these buttons, Nuix automatically runs the search and the items in the Results pane are updated.
- **Search** - free text field into which you can type or paste a search query. The Search field can contain millions of characters.
- **Run** - runs the search in the Search field.
- **Date filter** - offers four options that you can use in conjunction with the calendar controls: Between, Not Between, After, and Before. By default, searches are set to the option No date filter.
- **Calendar controls** - two calendar controls allow you to specify one or two dates in time to use in conjunction with the Date filter, including year, month, and day. Click the drop-down arrow to select a date using the visual calendar tool or type in the date you want to use in the field.
- **Clear** - clears the Search field and all filters, and sets all search criteria back to the default settings.
- **Advanced** - shows the Advanced Query Builder tool for building more complex queries without needing to know specific Nuix or Lucene search syntax.

For more information, see the Search section and further details about searching for items by date.

Using Keywords and Dates

Located at the top of the window, the Search bar gives you quick access to keyword and Nuix search syntax searches as well as options to search by date.

To perform a search from the Search bar:

1. Type directly into the Search field or cut and paste a predefined query into the field. The Search field can hold an unlimited number of characters; so, queries can be as long as necessary and use Boolean operators such as AND, OR, and NOT between search terms, and quotes around phrases. For more information on the supported search syntax, see the Nuix Search Guide.
2. If required, use the date filter to search between after or before certain dates, or use the Not between option to exclude a specific date range. The date filter searches on the Nuix-defined metadata filed called Item Date. For email messages, this is the Nuix Communications Date which is the Map-Client-Submit-Time, Sent Date, or Date metadata property. For loose files, this is the File Modified date or, if that is not present, it is the File Created date. If the item does not have any of these date fields, the
item date of the parent item is used. The left date control will search starting from 00:00:00.000 (HH:MM:SS.sss) and the right date control will search until 23:59:59.999 of the selected date. Click Search or press the Enter key to run the search.

Other actions you can perform are:

- View and reuse prior searches using the Back and Next arrow buttons in front of the Search field.
- Clear the search keywords and date filter by clicking Clear. When you clear a search, any selected nodes in the Filtered Items pane are cleared as well.

Note: Nuix allows for nested searches within the current results set allowing further searches to further narrow the results previously returned. The search bar should always be cleared between different searches to ensure they are being executed on the complete results set available ensuring that any future executions of the same search syntax produce the same results.

Searching with Filters

The Document Navigator pane on the left side allows you to narrow down your search results by selecting:

- Only the evidence that you want to locate by evidence folder or custodian. In the Evidence pane, clear the nodes you do not want included in the search results or in the Custodian pane, select whose data you do not want included.
- The types of items you want to locate by selecting the desired metadata filters in the Filtered Items pane.

All subsequent searches using the Search bar will exclude any items not selected in evidence, custodians and filtered items that are not selected panes in the Document Navigator.

Viewing Search Results

The Results pane displays a list of the items that match your selected criteria, whether from filters in the Document Navigator, a search, or other operations. By default, the view of the items is in a tabular list format, showing the metadata in columns from the associated metadata profile applied to that view.

At the bottom of the selected view, Nuix tells you how many items are displayed that match the criteria. If more items exist than can be shown in the view, it will state that. It also shows how many items are selected, if any, and how many items are removed from the list because they are immaterial or duplicates, if you use those options.

The Results pane comprises the following controls:

- **View By** - sets the view to show the items that match the selected criteria (that is, the items currently in the result set). Views include: Results, Thumbnails, Per-day Timeline, Timeline, Files, Words, Topics, Addresses, Event Map, Shingles, Entities and Network (some views only become visible if previously selected in the pre-processing options).
- **Hide immaterial items** - suppresses items that are not included in a legal export. Immaterial items are those items that are extracted for forensic completeness, but do not necessarily have intrinsic value in a legal context. Additionally, these items will NOT be exported as part of a legal export and are not included in the total size calculation for audited licenses. These items include, folders (file system, email, etc.), embedded inline graphics (email signatures, text items in PDF files, embedded objects without text, the zip container itself (not the contents), and mailbox files (PST, OST, NSF, MBOX, etc.).
- **Deduplicate results** - filters the items in the result set by MD5 hash to show only one of an item if it has duplicates. It is also possible to deduplicate at custodian level, this means that the deduplication will only remove duplicate items within the same custodian's data. Selecting this option increases the amount of time it takes to load a view. Click None to view items without deduplication.
- **View area** - displays the items or data in the format of the view you selected in the View by control. The columns in the default Results view can be changed by right-clicking on a column header and choosing from one of the available options.
- Add Tags - opens the Add Tags dialog so that you can apply tags to the selected items. This button is enabled when you select items in the result set in the Results, Thumbnails, and Addresses views.
- Exclude Items - opens the Exclude Items dialog so that you can exclude the selected items. This button is enabled when you select items in the Results, Thumbnails, and Addresses views.
- Export - allows you to select from a variety of export options and opens the corresponding dialog. Export options include exporting by view, items, case subset, annotations, and digest lists.

By default, the Results pane lists all items returned by a given query (search, filter, etc.).

Each row within the Results view is an active link and double-clicking a row displays the item in the Preview pane. You can customize the columns displayed in the result set with a metadata profile. The result set defaults to display 1,000,000 items. You can configure this value from File > Global Options > Viewer Limits.

You can change the view in the Results pane to display and interact with the data in different ways. The following topics explain how to interact with the various views using the available controls.

Nuix supports email threading to group email messages together so that legal teams can review them and make bulk decisions quickly. It presents email discussions with thread indexes and endpoint-inclusive status so reviewers can understand a whole conversation by reading just a few messages, rather than multiple iterations of the same content. This typically reduces the number of items for review by two-thirds.

Results View: Columns

You can customize the Results view’s columns in several ways by right-clicking on the column header and selecting a command.

Selecting the column header with the mouse sorts the result set. If the result set is very large, this can cause the application to appear hung. Repeatedly selecting the column causes Nuix to cycle through the types of sorting options, further delaying a responsive interface.

The following lists the right-click commands for columns:

- **Choose Column Profile** - lists the metadata profiles that you can use to change the metadata values that are displayed in the Results table view.
- **Metadata Name Column: Sort Ascending** - sorts the items in the column, starting with items that start with special characters, followed by items that start with numbers beginning with zero, and lastly items in alphabetical order beginning with the letter A.
- **Metadata Name Column: Sort Descending** - sorts the items in the column, starting with items in reverse alphabetical order that start with the letter Z, followed by items that start with numbers beginning with the highest number first, and lastly items with special characters in reverse order.
- **Metadata Name Column: Compute Distinct Values** - finds and displays all of the unique values in a given column. Each row is a unique record, and no parsing is performed within any field. The results of the Distinct Values calculation can be copied and pasted as comma separated values.
- **Metadata Name Column: Compute Column Sum** - totals all of the numerical values in a given column. Primarily for use with metadata whose values range in size, such as Digest Input Size. This dialog box displays selectable text.
- **Reset Sort Order** - resets the column to the Nuix default sort order, which is the order in which the documents were displayed when the search or filter operation was performed.

Results View: Rows

You can customize the data in the Results view’s rows in several ways, by selecting or clicking on items and with right-click operations.

In the Results table, you can perform these actions:

- To highlight a single row and display the item in the Preview pane, single-click on the row with the mouse or use the up or down arrows on your keyboard.
To select one or more highlighted items in the list, press the space bar.
To select all visible rows in the Results view, select the check box at the top of the table, or use Ctrl + A on the keyboard to select all visible rows in the Results, Word List, Statistics, History, and Thumbnail views.
To clear all visible rows in the Results view, clear (deselect) the checkbox at the top of the table, or use Ctrl + Shift + A on the keyboard to clear all visible rows in the Results, Word List, Statistics, History, and Thumbnail views.
To highlight contiguous rows of items, single-click an item and drag the mouse down or up to select additional rows or select the first item and press the Shift key and then select the last item to select all rows in between.
To highlight non-contiguous rows, select the first item and press the Shift + Ctrl keys while selecting additional rows. A right-click on any row or rows displays a context-sensitive set of commands. Some commands are only available if the item is selected (that is, the checkbox on that row is selected).

The following lists the right-click commands for rows:

- **Copy** - copies the selected rows to the clipboard. Includes just the metadata displayed by the current metadata profile.
- **Copy Cell Value** - copies the value of the selected cell to the clipboard.
- **Select All** - selects all visible rows in the Results view.
- **Select None** - clears all visible rows in the Results view.
- **Export** - exports items in a variety of ways using the Export controls.
- **Tags** - adds tags to and removes tags from selected items, including to items in the associated family and/or duplicates.
- **Custom Metadata** - adds and removes custom metadata, and applies template.
- **Custodian** - adds the selected items to a new or existing custodian, or removes items assigned to a custodian with options to include associated family items.
- **Item Set** - adds selected items to a new item set or an existing item set, and removes the selected items from an item set.
- **Review Job** - adds the selected items to or removes selected items from an existing Fast Review job including items in the associated family.
- **Production Set** - enables various operations on production sets including:
  - Create production sets
  - Add items to a new or existing production set
  - Re-number the Production Set items
  - Generate print previews for items prior to export
  - Remove items from the production set
  - Delete existing print previews for items
- **Cluster Run** - provides options to generate groups of chained near-duplicate clusters an item belongs to and view related email threading, add items to an existing cluster, and, remove clusters.
- **Automatic Classifier** - provides options to create an automatic classifier; build, export, and import models; add training items; automatically classify items.
- **Reload Items from Source Data** - reloads selected items from source data.
- **Scan for new Child Items** - scans new child items and processes only new items found to place them into the accurate location within the data tree.
- **Carve Unidentified Items** - carves the selected unidentified item. This can be viewed in the Filtered Items tree.
- **Exclude Items** - excludes items from being available for further case activity. This suppresses the items within the data set, including items in the associated family and/or duplicates.
- **Populate Stores** - allows the regeneration of both the binary natives store and the PDF image store with options to format the PDF images on generation.
- **Slip Sheet** - allows a slip sheet to be generated per container item and the resulting PDF included into the OCR process. A slip sheet is a placeholder for items that were not rendered correctly. This feature is not available for Proof Finder license.
- **Sample Items** - chooses a sample from the selected items and displays the sample in a new Workbench tab.
- **Perform OCR** - performs OCR and inserts searchable text and PDFs back into Nuix. This option is available only if ABBYY is installed. This feature is not available for Proof Finder license.
- **Bulk Redaction** - processes bulk redaction by selecting the word list or named entity group, creating a new markup set or using an existing markup set, and specifying the required PDF settings for the redaction.
- **Pivot** - allows you to view events that occurred before or after the selected item was created or by location.
- **Show** - finds items for the following:
  - **Show All Children Metadata** - generates a new Workbench containing a dynamically generated metadata profile for items contained in certain file types such as registry files and SQLite databases.
  - **Show All Descendants** * - finds all child items for the selected items.
  - **Show All Top-level Items** *: finds the highest-level ancestors for the selected items.
  - **Show All Families** - finds the highest-level ancestors and all child items for the selected items, including the items themselves with the results.
  - **Show All Near Duplicates** * - finds all items considered to be near duplicates of the selected items.
  - **Show All Chained Near-Duplicates** * - finds all items that are considered near duplicates of the selected items as well as all near duplicates of those new items in a new Workbench tab.
* This result set does not include the items themselves.

**Thumbnails View**

In the **Results** pane, you can view thumbnails of the images in that result set by selecting **View by: Thumbnails**.

If you use the **Skin Toned Image** filter in conjunction with this view, you can review images based on degrees of skin tone.

The thumbnails view indicates items with tags and video files.

In the Thumbnails view, you can perform the following operations:

- Click on an item to view the item in the Preview pane.
- Use the Zoom bar to zoom in and out of the Thumbnail grid, adjusting the display size of the Thumbnails. To zoom into an individual item, click on the item to view and use the zoom option in the Preview Pane.
- Use the Blur Images option to display all images in the Thumbnails grid as blurred.
- Select the item by clicking the check box below the image and right-click on the item to display a context-sensitive set of commands.

The following lists the right-click commands for items in the Thumbnails view.

- **Copy** - copies the selected item to the clipboard.
- **Select All** - selects all items in the Thumbnails view.
- **Select None** - clears all items in the Thumbnails view.
- **Export** - exports items in a variety of ways using the Export controls.
- **Add Tags** - adds tags to selected items.
- **Remove Tags** - removes tags from the selected items.
- **Add to Review Job** - adds the selected items to an existing Fast Review job.
- **Remove from Review Job** - removes the selected items from an existing Fast Review job.
- **Exclude Items** - excludes items from being available for further case activity. This suppresses the items within the data set, including items in the associated family and/or duplicates.
- **Reload Items from Source Data** - reloads selected items from source data.
- **Assign Custodian** - adds the selected items to a new or existing custodian with options to include associated family items.
- **Unassign Custodian** - removes the selected items from the selected custodian with options to include associated family items.
- ***Show All Descendants** - finds all child items for the selected items.
- ***Show All Top-level Items** - finds the highest-level ancestors for the selected items.
- **Show All Families** - finds the highest-level ancestors and all child items for the selected items, including the items themselves with the results.

* This result set does not include the items themselves.

### Per-Day Timeline View

In the **Results** pane, you can view items in your result set in a timeline by selecting **View by: Per-Day Timeline**. This view presents the number items with item date on each day. The size of the dot is directly proportional to the number of items; that is, bigger dots represent a higher number. You can view the date, Kind, Custodian, Tags, Languages and the Count of items created on that date. You can view the items grouped by:

- None
- Item Kind
- Tags
- Custodian
- Language

### Timeline View

In the **Results** pane, you can view items in your result set in a timeline by selecting **View by: Timeline**. This view presents a timeline based on each item. For each item, it displays the Date, Date Field, Name, Kind, Custodian, Languages and Tags. You can view the items grouped by:

- None
- Item Kind
- Tags
- Custodian
- Language

### Entities View

In the **Results** pane, you can view the entities related to Company, Credit Card, Email, Money, Personal ID, Country, IP address, and URL in that result set by selecting **View by: Entities**.

- **Company** - displays results related to company
- **Email** - displays results related to email addresses
- **Money** - displays results related to currencies
- **Country** - displays results related to country
- **IP Address** - displays results related to IP Addresses
- **URL** - displays results related to URLs

You can use the following functions in the Entities view:
- Show the entities found within the data set grouped by entities found.
- Filter the results by typing in the text to be matched in the entities, a particular company or card type, for a quick filter to narrow results.

**Files View**

In the **Results** pane, you can view the statistics related to just the items in that result set by selecting **View by:** Files.

The Files view provides information (statistics) about the number of processed and irregular files by file type within the current result set, as opposed to the Statistics tab that offers a look at file type statistics for the entire case.

In the Files view, you can perform the following operations:

- Re-sort the rows by toggling between ascending and descending order, click on a column header.
- View the items associated with one of the file types, double-click on the row to create a new Workbench tab displaying those items in a new result set.
- Export the view by selecting **Export.** For more information about exporting views, see Exporting Information from a View.

**Addresses View**

In the **Results** pane, you can view addresses in that result set by selecting **View by:** Addresses.

Select **Group by Domain** to view the results grouped by domain. It allows you to select options From, To, Cc, Bcc to view results. By default, the Addresses View uses the previous settings when executing a new search.

**Shingles View**

In the **Results** pane, you can view shingles in that result set by selecting **View by:** Shingles.

- View hits for just ASCII tokens, numerical tokens, non-ASCII tokens, Atypical Length tokens or all tokens.
- Use the filter to narrow down the returned results in the shingle list.

**Words View**

In the **Results** pane, you can view a list of all the words from the items in the current result set, as well as the frequency of their occurrence. The word list includes all words from both the content and properties of the selected items.

In the Words view, you can perform the following operations:

- To sort the rows, toggling between ascending and descending order, single-click on a column header.
- View hits for just ASCII words, numbers in the text, non-ASCII words, words that are an Atypical Length or all terms found.
- Word lists can be created from searching within the content of the documents selected or from just the properties of the selected documents.
- Use the filter to narrow down the returned results in the word list.
- To view the items that include a specific word in the list, double-click on the row to create a new Workbench tab displaying those items in a new result set.

**Topics View**

In the **Results** pane, you can view a list of all the topics from the items in the current result set, as well as their frequency. This is done after applying a list of stop words followed by stemming and frequency analysis. Nuix extracts the top five sentences that best represent the text in the selected items (documents). The top five sentences are selected based on a “frequency analysis” that determines the rank of a sentence. Sentences that
contain the most frequent words are ranked the highest. The view then presents the sentences ordered according to their original flow.

In the Topics view, you can perform the following operations:

- Re-sort the rows, toggling between ascending and descending order by clicking on a column header.
- View hits for all topic keywords, mostly ASCII topic keywords, mostly non-ASCII topic keywords, some numerical topic keywords, non-ASCII words, or topic keywords that are an Atypical Length.
- View the items that include a specific topic in the list, double-click on the row to open a new Workbench tab displaying those items in a new result set.
- Use the Topic Keyword Filter to narrow down the returned results in the topics list.
- Set the maximum number of topics to be displayed in the results list. Nuix calculates and displays this number as a percentage of the total number of items in your case. The default is 1%.

**Event Map View**

In the Results pane, you can view items in a specific thread to learn who was involved and to follow a conversation or document over time, by selecting View by: Event Map. The Event Map view provides a static graphical view of communications in the result set against a time line, showing who sent them and how they were sent to others.

In the Event Map view, you can change how the address labels in the right-hand column of the Event Map are displayed by selecting options in the drop-down Address menu:

- **None** - suppresses the display of the email address.
- **Personal** - displays only the personal portion of each email address. For example, Joanne Thomas would only display "Joanne Thomas".
- **Address** - displays only the address portion of each email address. For example, Joanne Thomas would only display "Joanne.C.Thomas @nuix.com".
- **Personal or Address** - displays either the Personal or Address portion of the email address depending on its availability.
- **Formatted Address** - displays the fully formatted email address. For example, Joanne Thomas would display "Joanne Thomas".

You can also export the Event Map view.

**Note:** All dates and time are stored as system time, which is essentially the number of ticks since 1970. When the items are then displayed, Nuix applies the appropriate time zone defined in the Case Properties dialog box, and presents the appropriate time.

**Network View**

In the Results pane, you can analyze patterns of communication between persons in a set of evidence by selecting View by: Network. The Network view provides a dynamic view of the communication patterns, including frequency of communication and any outlying communications in a graphical format.

You can control the format of the Network view by using the commands in the Networks menu.

You can filter the items that are displayed in the Network diagram by selecting the following options:

- **Direct (To)** - shows or hides items listed in the communications To field.
- **Indirect (Cc)** - shows or hides items listed in the communications Cc field.
- **Hidden (Bcc)** - shows or hides items listed in the communications Bcc field.
- **Show link count** - sets the minimum number of communications that must have occurred for items to display in the diagram. If the diagram is dense and you want to view fewer items based on frequency of communications, increase the value in this field.

The Network diagram is a dynamic view of the communications information associated with the specific result
set, meaning the display changes as you filter or change the result set. You can also customize the view in the following ways:

- **Run Layout** - freezes or unfreezes the automatic placement of the nodes in the diagram. When selected, the diagram is active and works to display the nodes in the most readable layout for viewing. The nodes will continue to pulse as the application continually optimizes the view. When you clear this option, the diagram is locked in place (although you can still manually move the nodes).
- **Scroll and pan** - Hold down the left mouse button on the background (white) area of the view to scroll the diagram up, down, left, or right. This does not change the rotation or proximity of the nodes.
- **Zoom** - Hold down the right mouse button on the background and move the cursor up or down on the Y axis of the screen to zoom in or out on the diagram.
- **Rotate** - Hold down the right mouse button on the background and move the cursor left or right along the X axis of the screen to rotate the diagram clockwise or counter-clockwise.
- **Highlight all communications partners** - Move the mouse over an address or left-click on an address to highlight all of the partners in the view who have communicated with that specific address.
- **Highlight two communications partners** - Move the mouse over a line or left-click on a line between two addresses to highlight the two addresses associated with that communication.
- **View items sent between two addresses** - Double-click on a communications line to display those items in a new result set view.

You can also export the Network view.

**Map View**

Investigators have large amounts of digital evidence involved in each case and need better ways to prioritize on the evidence that matters to the investigation. Using the Map view, investigators can now narrow down to the critical information faster, at less cost, using key location or times from investigation.

In the **Results** pane, you can view the map by **View by: Maps**. The Map view allows users to pin longitude and latitude float values from within a Nuix property field and plot the co-ordinates on a map view. Maps can be viewed as Cluster, Points or Heatmap. Users can either select to represent each individual item on the map using the points setting, or when dealing with larger quantity of items show areas of high activity using the heatmap setting.

For versions prior to Version 7.2, ensure the Nuix application has an active internet connection in order to use the map view. From Version 7.2 onwards, Map view can be viewed offline to access evidence from online sources such as Google Maps without leaving your investigative environment.

The map view is particularly useful to show where camera images were taken by plotting the co-ordinates from the exit information, or when resolving IP addresses found in log files to a geographical location and using the heatmap function to display locations from where high volumes of traffic originated.

With Pivot by Location and Time, operators can select an item, and, find all items with a timestamp within 30 seconds, 1, 5, 10, 20 minutes, 1 hour and 1 day. Similarly, they can select a file with GPS coordinates and pivot to find all items that have GPS coordinates within a specific radius. Map view allows you to perform search around a given point from Version 7.2 onwards.

**Hide Immaterial Items Option**

Nuix accounts for all items that it encounters during processing. This included system folders, folders inside a PST, views in an NSF, embedded objects in Office documents, inline graphics in email messages, and embedded items inside a PDF. Nuix does not provide a means of controlling how deep the extraction goes. Nuix extracts everything that it can find.

Nuix extracts and tracks all items, so that a complete and accurate accounting is performed, including maintaining a record of the entire evidence ancestry and all parent-child relationships.

However, in many contexts, the additional items that Nuix extracts are considered noise or immaterial to the goals of the investigator/reviewer/user. You can suppress these items by selecting **Hide in the Immaterial**
**items**: drop-down box which is found in the Results pane.

Immaterial items are those items that are extracted for forensic completeness, but do not necessarily have intrinsic value in a legal context. Additionally, these items are not exported as part of a legal export and are not included in the total size calculation for audited licenses.

Immaterial items include:

- Folders (file system, email, etc.)
- Embedded inline graphics (email signatures, embedded items in PDF files)
- Embedded objects without text
- The zip container itself (not the contents)
- Mailbox files like PST, OST, NSF, MBOX, EDB, STM, etc.

The best way to see exactly what has been determined an immaterial item is to directly query for them. From the Search field, type: `-flag:audited`

You can add all of the immaterial items to an exclusion set, and then view all the files that Nuix considers immaterial:

1. Across the entire case, search for **flag:audited**. This returns all of the immaterial items.
2. Select all items in the result set and then add them to an Excluded Items set called "Immaterial Items".
3. Clear the search results.
4. Run a search for **exclusion:Immaterial.Items**. This should return 0 hits.
5. From the Document Navigator, Excluded Items pane, uncheck the box next to the "Immaterial.Items" exclusion set.

You can now report on this at the statistics level as well as the result level by using a custom metadata profile to review the exact list of items that Nuix considers immaterial for a given collection of items.

**Deduplicate Results Option**

In the Preview pane, running a search with the Deduplicate results option selected only returns unique records to the result set. Removal of duplicates is based on the MD5 digest and is essentially performed by filtering on the MD5 digest field.

The copy that appears in the result set (the "original"), is the earliest item in the evidence tree as seen in the browser view. This ensures that each time a duplicate is removed the exact same item is always displayed/exported as part of the result set.

SHA-1 and SHA-256 hashes are only calculated for reference purposes. They are not used as part of the duplicate determination.

**Preview Pane**

The Preview pane, located on the Workbench tab, is comprised of information and tools that allow you to view the item itself, the metadata associated with the item and additional information to help analyze the context of the item. By default, the Preview pane is located on the right-hand side of the Workbench tab, and can be popped out of the window frame and/or resized within the Workbench window as necessary.

The Preview pane is comprised of three main areas:

- A toolbar at the top of the pane allows you to navigate between items, apply or edit comments, and view the item natively.
- An area with contextual information about the item, such as its source path and similar or related items.
- A set of tabs that present details about the item, such as the item's textual or image content and associated metadata.
Figure 12: Preview Pane

**Preview Toolbar**
A toolbar at the top of the Preview pane allows you to navigate between items, apply or edit comments, and view the item natively.

The toolbar is comprised of the following controls:

- **Previous Item** - Select the left arrow icon to preview the previous item in the result set.
- **Next Item** - Select the right arrow icon to preview the next item in the result set.
- **Item Name** - displays the Subject line of an email or the file name for all other item types.
- **Comment** - opens the Edit Comment dialog, allowing you to enter or edit a comment associated with the item being previewed. You can search for the text entered in a comment field.
- **Save As** - allows you to save the current item outside of the case. The default file type is selected depending on the item.
- **Launch** - opens the item in its native application if the application is installed on your system. When you create a case, selecting the option Store binary of data items will decrease the amount of time it takes to open an item natively.

**Preview Item Context**
Below the Preview pane toolbar is an area that provides contextual information about the item, such as the path where it existed in its source location, and other items similar or related to the item.

You can review the following information in this area:
- **Path** - shows the complete, hierarchical path that depicts all parent items for the item being previewed. You can view the items within the path by clicking any folder link, which opens a new results set.

- **Duplicates** - shows items that are exact and near duplicates of the item being previewed. Exact duplicate items are items with the same MD5 Hash value as the item being previewed. Near duplicate items are items that have a resemblance that is equal to or above the resemblance threshold set in Global Options.

- **Similar items** - shows items that are like the item being previewed. The High (90%+) similar), Medium (70%+) similar), and Low (50%+) similar) categories group like items by looking at the name of the item, the MD5 Hash value, and all words over six letters long that are the same.

- **Clusters** - shows the related clusters.

- **Related items** - shows the items that are a part of the same conversation thread as the item being previewed. An email thread is a series of emails that have been sent, forwarded, copied, and received, beginning with the first related communication. You can use the Event Map view to see who was involved with an email thread over time. Related items is only visible when dealing with email items.

**Preview Item Detail Tabs**

The **Preview** pane includes up to five Item Detail tabs.

![Figure 13: Preview Item Details](image)

These tabs present different views of the item’s content and associated metadata:

- **Email/Image/Text** - displays the extracted text of the item and details about the item, and is selected by default. Click **Details** to show or hide a subset of the metadata processed for the item, based on MIME type. Search words are highlighted within the preview. Use the dropdown on the top right to view a summary of the text, a selected number of lines or the full text. Your selection will be remembered when you view the next item. For an image, you can view the image on a transparent background by clicking Transparent Background, view full image or a thumbnail by selecting from the drop down menu, and zoom in and out of the image.

- **Family** - displays all attachments or child items associated with the item being previewed. Double-clicking an item in this tab opens a preview of the item in a new Workbench tab.

- **Metadata** - displays the metadata associated with the previewed item, including properties and Nuix-defined metadata. You can define the list of metadata that shows on this tab using a metadata profile; by default all metadata is shown. Click the drop-down menu and select a different profile to change the metadata shown. Custom Metadata: You can add custom metadata or apply a custom metadata template for the previewed item.

**Note:** The complete list of Nuix metadata fields is provided in Appendix A.

- **Add a custom metadata** - allows you to add a custom metadata. In the Add Custom Metadata dialog box:
Field name - allows you to specify field name.

Data type - allows you to select the type of data from the list- Integer, Float, Date, Text, Boolean. The data type defines the kind of data that can be entered in a metadata field.

Value - used to specify the value for the selected data type.

Edit custom metadata - allows you to select the custom metadata you wish to edit.

Remove custom metadata - allows you to select the custom metadata you wish to remove and click to remove the custom metadata. Confirm removal of the metadata by clicking Yes.

Apply custom metadata template - used to apply custom metadata template. In the Apply Custom Metadata Template dialog box:

Template name - Select the template you wish to apply, from the list. Note: If a field already exists, select an option from the list to retain the previous settings or overwrite them using this template.

Apply - allows you to select fields to apply the template.
  - All fields - used to apply the template to all fields
  - Selected fields only - used to apply only the selected fields. Specify the selected fields you wish to apply.

PDF - Renders a PDF view of the item. You can render PDFs while you are reviewing items, or incorporate it as part of the pre-review process. Note: Nuix uses the default settings to convert the file to PDF.

Nuix redacts sensitive content within the PDF by allowing you to create and edit markups, save markups, and highlight content. To add a new markup, click Create Markup Sets. The Markup Set dialog box is displayed. Specify the markup text to be added. Once the markups have been set up, click Save Markup Changes to apply them and save the changes to the PDF.

Change the layout of the PDF using the Rotate Left/Rotate Right button. You can change the page size by using the dropdown next to the rotate buttons. Click Zoom to view a pop-up menu to set the zoom level for the PDF.

Click Regenerate to regenerate the PDF from the native file replacing the currently stored PDF for that item.

To export the redacted information, navigate to Export > Legal Export and select the export option. Click Details to show or hide a subset of the PDF metadata and any text in the selected redaction. Click to replace the existing PDF of the item in the PDF store with a PDF generated outside of Nuix. Click to view the PDF in the native PDF viewer, if installed. At the bottom of the PDF view, are page controls that allow you to navigate multi-page PDF items.

Printed Image – renders a PDF view of the item.

Print Preview - displays a stamped version of the PDF.

Native - displays the current item within the preview pane using the application which is associated with the file type of the item.

Binary - displays the hex and allows you to search from the raw data structure at a binary level. This tab provides options that allows you to decode values recovered from unsupported file types and file fragments or unknown binary files. Select the required data decode format from the dropdown list at the bottom of the pane, and then select the first byte to interpret the complete date field.

Word List - displays a list of the words in the previewed item, along with how frequently each word appears in the item. You can begin typing in the Filter text field to go directly to a word or words matching those letters.

Entities - displays a list of all Named Entities found within the document being previewed. Click the Entities dropdown to view the different Entities found within the previewed document. You can begin typing in the Filter text field to go directly to a word or words matching those letters within the Entity list.

Person Named Entity - combines intelligent searching logic with lists of common first names,
surnames, salutations, and job titles. This enhances in areas such as eDiscovery and investigation to help identify unknown suspects or custodians. Person Named Entity is used to develop signatures for compliance with privacy regulations that prohibit storing different kinds of data, such as names and photos, that can be correlated.

- **Diff** - allows you to compare the current document with other similar documents to quickly identify the textual differences. Click **Set Item to Pivot** to set the current item as the pivot document to compare other documents against. Single click other documents in your results set to see their text appear in the left side of the Preview Pane with differences highlighted.

- **History** - displays the history showing the processing settings and any other user actions applied to the selected item. For example, added to Item Set, added to Production Set, added to Automatic Classifier. Each tab in the Preview pane presents horizontal or vertical scroll bars if the content does not fit in the viewing area.
Advanced Query Builder

Nuix offers an intuitive Advanced Search interface for building complex search queries. You can use the Advanced Query Builder to construct complex search expressions without knowing any search syntax. The Advanced Search tool builds the search syntax for you in the Search bar, and allows you to edit or remove parts of the expression to refine the query as you work.

Located on the far right of the Search, click Advanced to show the functionality. While you can type or paste large queries into the Search field, this tool allows those with limited knowledge of query syntax to build similar queries.

The Advanced Query Builder is comprised of the following controls:

- **Search Criterion Filter** - allows you to type the first letters of the search criterion for which you are looking and finds it in the list box.
- **Search Criterion List Box** - lists the types of criteria you can use to build a search expression. The associated options for each criterion are displayed in step two, to the right. You can use as many of these criteria as you wish in your query by adding them to the search expression one at a time.
- **Keywords: All of these words** - allows you to type in terms and phrases to use in the search in the associated free-text field on the right. The search returns only items that match all of the terms listed.
- **Keywords: Any of these words** - allows you to type in terms and phrases to use in the search in the associated free-text field on the right. The search returns items that match any of the terms listed.
- **Keywords: None of these words** - allows you to type in terms and phrases to use in the search in the associated free-text field on the right. The search returns only items that do not include the terms listed.
- **Keywords: Exact phrase** - allows you to type in an exact phrase to use in the search in the associated free-text field on the right. The search returns items that match only the exact phrase.
- **File size** - allows you to specify a minimum and maximum numerical file size to use in the search in the associated fields on the right. You must enter a value for both fields. File size is measured in bytes, and uses the Nuix Digest Input Size.
- **File type** - allows you to specify one or more file type(s) to search for in the associated list box on the right. Type into the filter control to go directly to a particular file type or file extension, or browse through the list of file types to find and select file extensions to include in your search expression. The file types you can choose from include application, audio, file system, image, message, server, text, and video. The list only includes the file extensions registered on the local system. You can select as many file types as you wish.
- **Tags** - allows you to select from the list of tags that exist in the case and match items that have the selected tag(s) applied to them. You can choose to match items with any, all, or none of the tags chosen with the drop-down control at the top of the list box.
- **Comments** - allows you to specify a text string in the associated free-text field on the right. The search returns only those items that include the string in the Nuix Comment field.
- **Custodians** - allows you to select a custodian from your list of custodians.
- **Item Sets** - allows you to select an item set from your list of item sets.
- **Production Sets** - allows you to select a production set from your list of production sets.
- **Document ID** - allows you to enter text that matches a Document ID.
- **Filters** - allows you to select a filter from the list.
- **Add to Expression** - adds the criteria you selected in steps one and two to the search expression, which is displayed in the Expression table. You must click this button each time you complete step two to add the expression to the query.
- **Expression Table** - displays each rule, or expression, as you add them. This collection of rules makes up the search query. You can choose to match all of any of the rules in the table, via the drop-down control at the top right of the table.
- **Edit** - allows you to select an expression in the table and edit that rule by loading the criteria you entered in steps one and two.
- **Remove** - removes the selected expression from the query.
- **Clear All** - clears all of the expressions from the Expression table.
- **Search** - runs the search.

You can hide or show the Advanced Query Builder by clicking **Advanced** at any time. The corresponding search syntax for the expressions you specified in the tool is displayed in the Search bar.

### Performing Advanced Searches

To create a search query using the Advanced Search query builder select **Advanced** in the Search bar to display the Advanced Search dialog box. Here you can build complex search expressions using the following steps:

1. Select **Pick search criterion** to select a criterion (option/ type of metadata) for your search. The available types are keywords, file size, file type, tags, comments, custodians, item sets, production sets, document ID, and filters.
2. Select **Search criterion options** to enter values for the selected criterion that require user-defined values. For example, if you select File size, specify the minimum and maximum range to match against.
3. Select **Add to Expression** to add the search syntax referred to as rules, to the query and is displayed in the Expression table.

Repeat steps 1-3 as needed until your query contains all the criteria/expressions you need for your search. Select whether to match all of the rules or any of the rules before you run the query. Select **Search**.

Other actions that can be performed in the Advanced Search query builder includes:

- **Edit** - allows you to edit the selected criterion within a rule in the expression table.
- **Remove** - allows you to delete a rule from the expression table.
- **Clear All** - allows you to clear the entire query from the expression table.
- **Advanced** in the search bar to close the Advanced Search query builder without searching.

Just like queries that you type into the Search bar, you can save the search queries built in this tool using the Saved Search option.

### Saving and Managing Search Queries

Nuix offers a wide variety of search syntax to refine search results. Refer to the Nuix Search Guide for details about the search syntax.

After you create a search query, you can save it for reuse at a future time.

#### Saving a Query

To save a search query:

1. Use the **Search** bar or **Advanced Search** window to create a search query.
2. Select the right arrow and select **Save Search**. The Save Search dialog box is displayed.
3. Give your search query a unique name and select the scope of the saved search. You may wish to use the search throughout the case, or by user, or local computer, or, the application provided. Select the option and click **OK**.

This search is now listed under the Saved Search for reuse.

When you save a search query, Nuix saves it in the following location `%AppData%\Nuix\Saved Searches`, in case you need to use a common set of search queries across multiple machines.

#### Loading a Saved Query

To load a saved query, click **Saved Search** menu, from the list of available searches, choose the query you
wish to rerun.

Nuix automatically runs the search when it is loaded. Any matching items are displayed in the Results list.

**More options on Saved Query**

To perform more actions on saved searches, open Saved Search in Global options. Click **File > Global Options > Saved Search**.

Select the saved searches you wish to perform the action on and:

- allows you to delete the selected saved searches. The selected query is removed from the following location, %AppData%\Nuix\Saved Searches, and is no longer available to load. However, it will remain in the case’s Search History list in the Document Navigator.

- allows you to rename a saved search.

**Move** - allows you to move the scope for the selected items. Select the scope from the drop-down menu.

**Duplicate** - allows you to duplicate the selected search query.

**Import** - allows you to import a saved search.

**Export** - allows you to export the selected saved search.

Nuix also saves all queries that you run as a history record within the History tab. To rerun a previously executed search from the History tab, double-click on the record with the search syntax desired. A new window will open with the returned results.

**Search and Tag Tab**

The Search and Tag tab allows you to search processed data using queries, and tag the results for future reference. You can search by individual keyword, proximity or freeform queries, by simply uploading a valid Search and Tag CSV or JSON file to run multiple search queries at once.

To open Search and Tag tab, go to **Windows > New Search and Tag Tab**.

Search and Tag activities use CSV Files, which are similar to an excel-type format, but instead uses the file extension .csv (comma delimited) or a JSON file, which is based on data pairing whereby a value name is paired with a value. Since the JSON file saves settings, it makes the search portable and provides consistent results throughout all installations.

The fields that are searched:

- In GUI, fields listed in **Global Options > Search**
- In Scripting API, the common set of 4 fields that are also used elsewhere in scripting searches

The Search and Tag tab consists of:

- Search and Tag Table
- Search and Tag Basic Controls
- Advanced Search - See **Advanced Query Builder** topic for more details.
- Search and Tag Settings
- Scoping Query - allows you to enter the scoping query within the query box.
- Search and Tag Actions

**Search and Tag Table**

The Search and Tag table consists of rows of queries with tags. See Search and Tag Basic Controls to add/delete, move and clear rows. Once you have added the row, double-click on the row's tag field to enter a
tag and the query field to enter the query. Click the right arrow to run the query, the Search and Tag Results pane is displayed with the number of hits. Your table can also contain blank rows to create separate sections, blank rows are ignored when the table is run.

Results are populated in real time during the search and tag operation. The depth of results shown depends on the search options configured. This view can be saved as a CSV file.

**Search and Tag Basic Controls**

The basic controls include adding and deleting rows to the table, moving an entry, and clearing the table.

- Click the plus sign to add a row to the bottom of the table. If a row is selected while adding, then the new row is added below the selected row.
- Click the minus sign to delete one or more selected rows from the table.
- Click the up arrow to move one or more selected rows up.
- Click the down arrow to move one or more selected rows down.
- Click **Clear** to clear all current rows from the table.

**Search and Tag Settings**

Search and Tag Settings displays the following settings:

- **Query Scope** - Select the following options to:
  - **Include Families** - In addition to direct matches, Nuix will report on indirect family item matches and indirect top-level item matches.
  - **Deduplicate families** - Nuix provides matches value based on a family level deduplication. This option is available only if the “Included Families” is selected.
  - **Omit excluded items**
  - **Omit immaterial family items**
  - **Unique items** - Unique items only matches tags to one row

- **Action** - Select the option from the dropdown list for items that match query:
  - **Add tag**
  - **Remove tag**
  - **No Action**
  - Select the **Create tags even if no items match query** option, if you want to create tags for queries that result with no item matches, this will create empty expanded tags as well. This option is enabled only when you have selected Add tag as your action.
  - Select the **Remove empty tags from case** option, if you want to completely remove empty tags from the case. This option is enabled only when you have selected Remove tag as your action.
  - Select the **Show tag name columns in results** option, if you want view tag name columns in the results view. This is a default selection.

- **Validation** - Select the options to:
  - **Allow duplicate tags** - allows multiple queries to result in the same tag being applied
  - **Allow duplicate queries** - allows the same query to be used to create multiple tags
  - Select an option from the list for missing tags.
    - **Fill with query** - fills the missing tag with the matching query
    - **Fill with row number** - fills the missing tag with the row number
  - **Show validation error** - displays a validation error for missing tags

- **Expand Tags** - Select this option to expand tags. The prefix and suffix match the family items and deduplication and are switched ON and OFF with these selection. The tags expand into direct matches, top-level matches, family matches, deduplicated top-level matches, deduplicated family matches, and unique matches by appending the prefix and suffix to each tag in each row of the table. “|” is used as the notation to create hierarchical tags (see figure above). By using the pipe you can assemble the tags into families of related tags. e.g. “tag 1|raw hits”, “tag 1|top-level hits”, or alternatively “raw hits|tag 1”, “top-level hits|tag 1”, or even “batch 1|tag 1|raw hits” by using both the prefix and suffix at the same time. The rows in this table are enabled or disabled depending on what is selected in the query scope.
section. All fields are optional.

**Scoping Query**

Scoping query limits the scope of all direct matches in the table rows. Note that it applies only to direct matches. Items that are dragged in as families and top-level items are not limited by the scoping query.

**Search and Tag Actions**

The Search and Tag Actions allow you to perform actions:

- **Import** - You can import 2 kinds of files:
  - *e* - This file makes search and tag more portable and provides consistent results throughout all installations. JSON files can have preferred settings. If there is no data in the table already, then the settings from the JSON file will be used (if present). However, if the table already contains data then its present settings will be retained and the settings from the JSON file will be ignored.
  - *CSV file* - This file format does not retain any settings.

Importing files while there is data already will append the new rows to the end of the table and keep the existing rows and settings intact. This helps to build the query table.

JSON files should be in the format shown below:

![Figure 14: JSON File Format](image)

CSV files should be in the following format of two columns:

- The first column contains the tag name and supports the use of nested tags via the pipe (‘|’) character
- The second column contains any valid Nuix query or a simple keyword
- No header row

- **Save** - To save a search and tag file, click **Save**. The Save dialog box is displayed, specify the location, file name and extension (CSV or JSON) and click **Save**.

**Search and Tag Results Pane**

On building the Search and Tag table, click the right arrow to run the query, the Search and Tag Results pane is displayed with the number of matches.

**Save Search and Tag Results**

To save search and tag results, click **Save** and browse to the desired location. Click **Save**.
View Tagged items

To view the tagged items, go to Filtered Items > Tagged. The Tagged items display a folder structure for each tag created, and sub-folders for expanded tags.
### Appendix A: Nuix Metadata Fields

Nuix-defined Metadata includes:

<table>
<thead>
<tr>
<th>METADATA</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audited</td>
<td>Whether an item was flagged as audited. isAudited()</td>
</tr>
<tr>
<td>Audited Size</td>
<td>The auditable size of the item (*) — only meaningful if the case was audited when it was loaded. getAuditedSize()</td>
</tr>
<tr>
<td>Automatic Classifications</td>
<td>Classifications for automatically classified items.</td>
</tr>
<tr>
<td>Automatic Classifications (Selected)</td>
<td>Classifications for automatically classified items, for the Automatic Classifier(s) currently selected in the Document Navigator.</td>
</tr>
<tr>
<td>Automatic Classifier Confidence</td>
<td>The confidence (or probability of correctness) for automatically classified items.</td>
</tr>
<tr>
<td>Automatic Classifier Confidence (Selected)</td>
<td>The confidence (or probability of correctness) for automatically classified items, for the Automatic Classifier(s) currently selected in the Document Navigator.</td>
</tr>
<tr>
<td>Automatic Classifier Gain Confidence</td>
<td>The confidence (or probability of correctness) for automatically classified items as would be shown in a gain chart. If the predicted classification is the positive classification, this is identical to Automatic Classifier Confidence, otherwise it is 1.0 - Automatic Classifier Confidence.</td>
</tr>
<tr>
<td>Automatic Classifier Gain Confidence (Selected)</td>
<td>The confidence (or probability of correctness) for automatically classified items as would be shown in a gain chart, for the Automatic Classifier(s) currently selected in the Document Navigator. If the predicted classification is the positive classification, this is identical to Automatic Classifier Confidence (Selected), otherwise it is 1.0 Automatic Classifier Confidence (Selected).</td>
</tr>
<tr>
<td>Bad Extension</td>
<td>Whether the file appears to have an irregular extension. Plain text and items without a file size property are excluded.</td>
</tr>
<tr>
<td>Batch Load GUID</td>
<td>The GUID for the batch load associated with this item if available.</td>
</tr>
<tr>
<td>Bcc</td>
<td>The hidden recipients (Bcc) for the communication, if the item was a communication. getCommunication().getBcc()</td>
</tr>
<tr>
<td>Binary Stored</td>
<td>Indicates whether the binary is stored in the database for this item. isBinaryStored()</td>
</tr>
<tr>
<td>Carved</td>
<td>Indicates that the item was carved out of slack-space or from unidentified item data.</td>
</tr>
<tr>
<td>Cc</td>
<td>The indirect recipients (Cc) for the communication, if the item was a communication. getCommunication().getCc()</td>
</tr>
<tr>
<td>Function</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Chained Near-Duplicate Count(**)</td>
<td>The number of chained near-duplicate items (does not include the item itself.)&lt;br&gt;<code>getChainedNearDuplicates(float).size()</code></td>
</tr>
<tr>
<td>Chained Near-Duplicate Custodian Set</td>
<td>The set of custodian names associated with this item and its chained near-duplicate items.&lt;br&gt;<code>getChainedNearDuplicateCustodianSet(float)</code></td>
</tr>
<tr>
<td>Chained Near-Duplicate GUIDs (**)*</td>
<td>A list of GUIDs of chained near-duplicate items (does not include the item itself.)&lt;br&gt;<code>getChainedNearDuplicates(float), then getGuid() for each child.</code></td>
</tr>
<tr>
<td>Chained Near-Duplicate Paths(**)</td>
<td>A list of paths to chained near-duplicate items (does not include the item itself.)&lt;br&gt;<code>getChainedNearDuplicates(float), then getPath() for each child.</code></td>
</tr>
<tr>
<td>Child Count</td>
<td>The number of children this item has.&lt;br&gt;<code>getChildren().size()</code></td>
</tr>
<tr>
<td>Child Material Count</td>
<td>The number of material children this item has.&lt;br&gt;<code>getChildren(), then count isAudited() children.</code></td>
</tr>
<tr>
<td>Child Names(**)</td>
<td>A list of names of child items.&lt;br&gt;<code>getChildren(), then getName() for each child.</code></td>
</tr>
<tr>
<td>Cluster Branch IDs</td>
<td>The list of string values indicating which threading branch an item is associated with in all the clusters it belongs to. Values take the form &quot;x.y&quot; where integer x is a cluster number, and integer y is a branch ID. Values are sorted according to cluster run, oldest first. An empty value implies there was insufficient threading information in a particular cluster to compute a value.&lt;br&gt;<code>getClusterBranchIds().values()</code></td>
</tr>
<tr>
<td>Cluster Branch IDs (Selected)</td>
<td>The list of string values indicating which threading branch an item is associated with in the currently selected clusters. Values take the form &quot;x.y&quot; where integer x is a cluster number, and integer y is a branch ID. Values are sorted according to cluster run, oldest first.</td>
</tr>
<tr>
<td>Cluster Endpoint Status</td>
<td>A list of string values indicating the item's endpoint in all the clusters it belongs to. Values are sorted according to cluster run, oldest first. An empty value implies there was insufficient threading information in a particular cluster to compute a value.&lt;br&gt;<code>getClusterEndpointStatus().values()</code></td>
</tr>
<tr>
<td>Cluster Endpoint Status (Selected)</td>
<td>The list of string values indicating the item's endpoint status in the currently selected clusters. Values are sorted according to cluster run, oldest first.</td>
</tr>
<tr>
<td>Cluster IDs</td>
<td>The list of chained near-duplicate clusters an item belongs to. Clusters are denoted by their run label followed by an integer cluster label. (e.g. myClusterRun-123.) Labels are sorted according to cluster run, oldest first.&lt;br&gt;<code>getClusterPivots().keySet()</code></td>
</tr>
<tr>
<td>Cluster IDs (Selected)</td>
<td>The list of chained near-duplicate clusters an item belongs to in the set of clusters currently selected in the Filtered Items pane of the Document Navigator.</td>
</tr>
<tr>
<td>Cluster Pivot Resemblances</td>
<td>A list of resemblance values, one for each of the clusters an item belongs to. Values are the resemblance values between the item and each cluster's pivot item. The list is sorted according to cluster run, oldest first.&lt;br&gt;<code>getClusterPivotResemblances().values()</code></td>
</tr>
<tr>
<td>Cluster Pivot Resemblances (Selected)</td>
<td>A list of resemblance values, one for each of the clusters an item belongs to in the set of clusters currently selected in the Filtered Items pane of the Document Navigator.</td>
</tr>
<tr>
<td>Field</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cluster Pivots</td>
<td>A list of boolean values, one for each of the clusters an item belongs to. A value of true indicates an item is the pivot member of the cluster it is contained in. A value of false means it is not a pivot item. The list is sorted according to cluster run, oldest first.</td>
</tr>
<tr>
<td>Cluster Pivots (Selected)</td>
<td>A list of boolean values, one for each of the clusters an item belongs to in the set of clusters currently selected in the Filtered Items pane of the Document Navigator.</td>
</tr>
<tr>
<td>Cluster Thread Indexes</td>
<td>A list of string values indicating the thread index an item is associated with in all the clusters it belongs to. Values are sorted according to cluster run, oldest first. An empty value implies there was insufficient threading information in a particular cluster to compute a value.</td>
</tr>
<tr>
<td>Cluster Thread Indexes (Selected)</td>
<td>A list of string values indicating the thread index of an item in the currently selected clusters. Values are sorted according to cluster run, oldest first.</td>
</tr>
<tr>
<td>Comment</td>
<td>The comment added to the item by users.</td>
</tr>
<tr>
<td>Communication Date</td>
<td>The date of the communication, if the item was a communication (e.g. an email.)</td>
</tr>
<tr>
<td>Custodian</td>
<td>The custodian assigned to the item.</td>
</tr>
<tr>
<td>Decrypted</td>
<td>Indicates the item has been decrypted.</td>
</tr>
<tr>
<td>Deleted</td>
<td>A flag indicating that the item was deleted, or one of its ancestors were deleted.</td>
</tr>
<tr>
<td>Deleted File - All Blocks Available</td>
<td>Indicates that the item is a deleted file that doesn't have any of its block allocated to live files. It may be fully recovered from unallocated space, or part of the file may have been overwritten at some point.</td>
</tr>
<tr>
<td>Deleted File - Metadata Recovered</td>
<td>Indicates that the item is a deleted file that has had its metadata recovered from unallocated space. Either all of the file was overwritten with other data or the file record couldn't be linked back to its data.</td>
</tr>
<tr>
<td>Deleted File - Some Blocks Available</td>
<td>Indicates that the item is a deleted file that has some of its block allocated to live files. Only part of the data will be recovered from unallocated space since some areas were overwritten with other data.</td>
</tr>
<tr>
<td>Digest Input Size</td>
<td>The number of bytes which ran through the digest. This can be used as an approximate indicator of the size of the item, or as an additional sanity check for comparing two items’ digests.</td>
</tr>
<tr>
<td>Document IDs</td>
<td>All document IDs that have been assigned to the item.</td>
</tr>
<tr>
<td>Document IDs (Selected)</td>
<td>Document IDs that have been assigned to the item in the currently selected production sets.</td>
</tr>
<tr>
<td>Duplicate Count(***</td>
<td>The number of duplicate items (does not include the item itself.)</td>
</tr>
</tbody>
</table>
| **Duplicate Custodian Set** | The set of custodian names associated with this item and its duplicate items.  
getDuplicateCustodianSet() |
| --- | --- |
| **Duplicate GUIDs (**)** | A list of GUIDs of duplicate items (does not include the item itself.)  
getDuplicates(), then getGuid() for each child. |
| **Duplicate Item Dates (**)** | A list of item dates of duplicate items (does not include the item itself.)  
getDuplicates(), then getItemDate() for each child. |
| **Duplicate Paths(***) | A list of paths to duplicate items (does not include the item itself.)  
getDuplicates(), then getPath() for each child. |
| **Encrypted** | A flag indicating that the item was encrypted.  
isEncrypted() |
| **Entity: Company** | The company named entities identified in the items text.  
getEntities(entityType) |
| **Entity: Country** | The country named entities identified in the items text.  
getEntities(entityType) |
| **Entity: Credit Card Number** | The credit card number named entities identified in the items text.  
getEntities(entityType) |
| **Entity: Email** | The email named entities identified in the items text.  
getEntities(entityType) |
| **Entity: IP Address** | The IP address named entities identified in the items text.  
getEntities(entityType) |
| **Entity: Money** | The money named entities identified in the items text.  
getEntities(entityType) |
| **Entity: Person** | The person named entities identified in the items text.  
getEntities(entityType) |
| **Entity: Personal ID** | The personal ID named entities identified in the items text.  
getEntities(entityType) |
| **Entity: URL** | The URL named entities identified in the items text.  
getEntities(entityType) |
| **Exclusion** | The exclusion reason added to the item by users.  
getExclusion() |
<p>| <strong>Failed</strong> | Indicates the item has a failure detail property. That is, an error occurred processing the data. |
| <strong>Family Inline</strong> | Indicates the item and its family members are present in the one evidence database. This flag was not present in v3.6 cases and earlier. |
| <strong>File Data</strong> | Indicates the item is a readable file. |</p>
<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Extension (Corrected)</td>
<td>Holds a file extension which is consistent with the File Type. If the source data had an inconsistent file extension then this will be different to the original file extension.</td>
</tr>
<tr>
<td></td>
<td>getCorrectedExtension()</td>
</tr>
<tr>
<td>File Extension (Original)</td>
<td>Holds the file extension as it was found in the original source data.</td>
</tr>
<tr>
<td></td>
<td>getOriginalExtension()</td>
</tr>
<tr>
<td>File Size</td>
<td>The file size of the item, or null if not a file.</td>
</tr>
<tr>
<td></td>
<td>getFileSize()</td>
</tr>
<tr>
<td>File Type</td>
<td>The file type of the item. These resemble MIME types but many of the types were invented by us as there was no registered type for their respective file formats.</td>
</tr>
<tr>
<td></td>
<td>getMimeType()</td>
</tr>
<tr>
<td>From</td>
<td>The senders (From) for the communication. Generally there will be only one, but it is possible for there to be more than one.</td>
</tr>
<tr>
<td></td>
<td>getCommunication().getFrom()</td>
</tr>
<tr>
<td>GUID</td>
<td>A globally unique ID generated by Nuix which is unique for every item across all cases.</td>
</tr>
<tr>
<td></td>
<td>getGuid()</td>
</tr>
<tr>
<td>Hidden Stream</td>
<td>Indicates that the item is a hidden stream associated with another item. Examples of this include NTFS alternate data streams and HFS+ resource forks.</td>
</tr>
<tr>
<td>Identification Disabled</td>
<td>Indicates that file-type identification was not enabled when this item was processed.</td>
</tr>
<tr>
<td>Images Not Processed</td>
<td>Indicates the item’s images were not processed.</td>
</tr>
<tr>
<td>Inlined</td>
<td>If given item is being displayed as a part of outer item. e.g., an image in RTF document. Legal export, for example, excludes such items.</td>
</tr>
<tr>
<td>Irregular Item</td>
<td>Indicates whether the item was corrupted, deleted or encrypted.</td>
</tr>
<tr>
<td>Item Category</td>
<td>Either &quot;Email&quot;, &quot;Attachment&quot;, &quot;Electronic File&quot;, &quot;Electronic Directory&quot; or blank, depending on the file type and the position of the item in the tree.</td>
</tr>
<tr>
<td></td>
<td>getItemCategory()</td>
</tr>
<tr>
<td>Item Date</td>
<td>The date of the item. Derived from the communication date, or if the communication date is not present, the last modification date, or finally the creation date.</td>
</tr>
<tr>
<td></td>
<td>getItemDate()</td>
</tr>
<tr>
<td>Item Set Duplicate Count(*)</td>
<td>The number of duplicate items for an item in a selected item set (not including the original item).</td>
</tr>
<tr>
<td></td>
<td>ItemSet.findDuplicates(Item).size()</td>
</tr>
<tr>
<td>Item Set Duplicate Custodian Set</td>
<td>The set of custodian names associated with this item and its duplicate items for a selected item set.</td>
</tr>
<tr>
<td></td>
<td>ItemSet.findDuplicates(Item), then getCustodian() for each child.</td>
</tr>
<tr>
<td>Item Set Duplicate GUIDs (*)</td>
<td>A list of GUIDs of duplicate items for an item in a selected item set (not including the original item).</td>
</tr>
<tr>
<td></td>
<td>findDuplicates(), then getGuid() for each child.</td>
</tr>
<tr>
<td>Item Set Duplicate Item Dates (*)</td>
<td>A list of item dates of duplicate items for an item in a selected item set (not including the original item).</td>
</tr>
<tr>
<td></td>
<td>findDuplicates(), then getItemDate() for each child.</td>
</tr>
<tr>
<td>Item Set Duplicate Paths(*)</td>
<td>A list of paths to duplicate items for an item in a selected item set (not including the original item).</td>
</tr>
<tr>
<td></td>
<td>findDuplicates(), then getPath() for each child.</td>
</tr>
<tr>
<td>Item Sets As Duplicate</td>
<td>The item sets that this item is a member of as a duplicate. getItemSetAsDuplicate()</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Item Sets As Duplicate(**)</td>
<td>A list of item sets that this item is a member of as a duplicate. getItemSetsAsDuplicate().</td>
</tr>
<tr>
<td>Item Sets As Original</td>
<td>The item sets that this item is a member of as an original. getItemSetAsOriginal()</td>
</tr>
<tr>
<td>Item Sets As Original(**)</td>
<td>A list of item sets that this item is a member of as an original. getItemSetsAsOriginal().</td>
</tr>
<tr>
<td>Item Sets (**</td>
<td>A list of item sets that this item is a member of. getItemSets().</td>
</tr>
<tr>
<td>Kind</td>
<td>The file type kind of the item, such as &quot;Email&quot;, &quot;Containers&quot;, &quot;Spreadsheets&quot;, e.t.c. getItemKind()</td>
</tr>
<tr>
<td>Language</td>
<td>The detected language of the text of the item. getLanguage()</td>
</tr>
<tr>
<td>Licence Restricted</td>
<td>An item not processed due to licence restrictions.</td>
</tr>
<tr>
<td>Loose File</td>
<td>Loose files are the files which you would generally see in a file browser when browsing a directory. However, the files inside disk image and logical image formats are treated as the loose files instead of the outer image file. isLooseFile()</td>
</tr>
<tr>
<td>Markup Sets</td>
<td>All markup sets the item is included in. getPrintedImage() PrintedImage.getPages() Case.getMarkupSets()</td>
</tr>
<tr>
<td>Material Child Names (**</td>
<td>A list of names of material child items. getChildren(), then check isAudited(), then call getName() for each child.</td>
</tr>
<tr>
<td>MD5 Digest</td>
<td>Digests (hashes) for the item. Which are present will depend on the settings specified at load time, as well as the size of the data item. getDigests().getMd5()</td>
</tr>
<tr>
<td>MD5 Digest (Latest)</td>
<td>Latest digests (hashes) for the item. Which are present will depend on the settings specified at load time, as well as the size of the data item, and whether the item has been reloaded/replaced. getLatestDigests().getMd5()</td>
</tr>
<tr>
<td>MD5 Digest (Original)</td>
<td>Original digests (hashes) for the item. Which are present will depend on the settings specified at load time, as well as the size of the data item, and whether the item has been reloaded/replaced. getOriginalDigests().getMd5()</td>
</tr>
<tr>
<td>MIME Type</td>
<td>The file type of the item. These resemble MIME types but many of the types were invented by us as there was no registered type for their respective file formats. getMimeType()</td>
</tr>
<tr>
<td>Name</td>
<td>The name of the item. The meaning of this is generally dependent on the type of item, but the name is usually chosen to be a user readable value which you might expect to see if looking at the item in a native viewer for the format in which it is contained. getName()</td>
</tr>
</tbody>
</table>
| **Near-Duplicate Count (**)** | The number of near-duplicate items (does not include the item itself.)  
getNearDuplicates(float).size() |
|-----------------------------|--------------------------------------------------------------------------------------------------|
| **Near-Duplicate Custodian Set** | The set of custodian names associated with this item and its near-duplicate items.  
getNearDuplicateCustodianSet(float) |
| **Near-Duplicate GUIDs (**)** | A list of GUIDs of near-duplicate items (does not include the item itself.)  
getNearDuplicates(float), then getGuid() for each child. |
| **Near-Duplicate Paths (**)** | A list of paths to near-duplicate items (does not include the item itself.)  
getNearDuplicates(float), then getPath() for each child. |
| **Not Failed** | Indicates the item does not have a failure detail property. |
| **Not File Data** | Indicates the item is not a readable file. |
| **Not Loose File** | Items not marked as being loose files.  
!isLooseFile() |
| **Not Physical File** | Items not marked as being physical files.  
!isPhysicalFile() |
| **Not Processed** | Indicates the item was identified but not processed any further. |
| **Not Top-level** | Indicates the item is not a top-level item.  
!isTopLevel() |

**NOTES:**

| **Parent GUID** | The GUID of the parent data item.  
getParent().getGuid() |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Partially Processed</strong></td>
<td>Indicates the item's children were only partially processed. Some children were explicitly skipped at the direction of the user.</td>
</tr>
</tbody>
</table>
| **Path Name** | The name of each item from the root evidence container. In the application this is generally separated by a slash ('/') or arrows.  
getPathNames() |
| **Physical File** | Physical files correspond to the highest items in the data tree which have binary, and typically correspond to those files which were used as input evidence to the case.  
isPhysicalFile() |
| **Poisoned** | Indicates the item caused a critical error during processing on several attempts. |
| **Position** | The position of the item within the tree. Will be unique for every item in the case, and sorting on it will sort in the same order as the evidence is displayed in the document navigator.  
getPosition() |
| **Preview Text** | Provides a preview of the item's extracted text from the start of the text. |
| **Printed Image Generation Method** | The method used to generate the printed image, which is typically a PDF, if one has been generated for this item.  
getPrintedImageInfo().wasPrinted(), etc. |
<table>
<thead>
<tr>
<th><strong>Printed Image Page Count</strong></th>
<th>The number of pages in the printed image, which is typically a PDF, if one has been generated for this item.</th>
</tr>
</thead>
<tbody>
<tr>
<td>getPrintedImageInfo().getPageCount()</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Production Sets</strong></th>
<th>All production sets the item has been assigned to.</th>
</tr>
</thead>
<tbody>
<tr>
<td>getProductionSets()</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Production Sets (Selected)</strong></th>
<th>Production sets the item has been assigned to in the currently selected production sets.</th>
</tr>
</thead>
<tbody>
<tr>
<td>getProductionSets()</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Reloaded</strong></th>
<th>Indicates the item was reload into Nuix.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>SHA-1 Digest</strong></th>
<th>getDigests().getSha1()</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>SHA-1 Digest (Latest)</strong></th>
<th>getLatestDigests().getSha1()</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>SHA-1 Digest (Original)</strong></th>
<th>getOriginalDigests().getSha1()</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>SHA-256 Digest</strong></th>
<th>getDigests().getSha256()</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>SHA-256 Digest (Latest)</strong></th>
<th>getLatestDigests().getSha256()</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>SHA-256 Digest (Original)</strong></th>
<th>getOriginalDigests().getSha256()</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Shannon Entropy</strong></th>
<th>The Shannon Entropy calculated for the item based off its binary data. The value will range from 0.0 to 8.0. A value of 0.0 indicates completely uniform data (i.e. all one value), while a value of 8.0 is perfectly random data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>getShannonEntropy()</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Skin-tone</strong></th>
<th>The confidence score for skin-tone images, ranging from 0.0 (low confidence) to 1.0 (high confidence).</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Slack Space Region</strong></th>
<th>Indicates that the item represents a region of recovered slack-space.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>SSDeep Fuzzy Hash</strong></th>
<th>The SSDeep fuzzy hash string calculated for the item based on its binary data. The value takes the form blocksize:hash1:hash2 where blocksize is an integer, and the subsequent hashes are base-64 encoded.</th>
</tr>
</thead>
<tbody>
<tr>
<td>getSSDeepFuzzyHash()</td>
<td></td>
</tr>
<tr>
<td><strong>Thread Count</strong>(***)</td>
<td>The number of items in the same discussion thread (does not include the item itself.)&lt;br&gt;getThreadItems().size()</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Thread GUIDs</strong>(***)</td>
<td>A list of GUIDs of items in the same discussion thread (does not include the item itself.)&lt;br&gt;getThreadItems(), then getGuid() for each child.</td>
</tr>
<tr>
<td><strong>Thread Paths</strong>(***)</td>
<td>A list of paths to items in the same discussion thread (does not include the item itself.)&lt;br&gt;getThreadItems(), then getPath() for each child.</td>
</tr>
<tr>
<td><strong>To</strong></td>
<td>The direct recipients (To) for the communication, if the item was a communication.&lt;br&gt;getCommunication().getTo()</td>
</tr>
<tr>
<td><strong>Top-level</strong></td>
<td>Indicates the item is considered a top-level item, since all of its ancestor items are containers. e.g. Loose files that are not containers, Office documents inside a zip container, emails inside mailboxes etc.&lt;br&gt;isTopLevel()</td>
</tr>
<tr>
<td><strong>Top-level GUID</strong></td>
<td>The GUID of the top-level data item.&lt;br&gt;getTopLevelItem().getGuid()</td>
</tr>
<tr>
<td><strong>Top-level Item Date</strong></td>
<td>The Item Date of the top-level data item.&lt;br&gt;getTopLevelItemDate()</td>
</tr>
<tr>
<td><strong>Top-level Path Name</strong></td>
<td>The name of each item from the root evidence container to the top-level item. In the application this is generally separated by a slash (’/’) or arrows.&lt;br&gt;getTopLevelItem().getPathNames()</td>
</tr>
<tr>
<td><strong>Training Classifications</strong></td>
<td>Classifications for items provided as training data for automatic classifiers.</td>
</tr>
<tr>
<td><strong>Training Classifications (Selected)</strong></td>
<td>Classifications for items provided as training data for the automatic classifiers(s) currently selected in the Document Navigator.</td>
</tr>
<tr>
<td><strong>Unallocated Space</strong></td>
<td>Indicates that the item represents a region of recovered unallocated space in the file system.</td>
</tr>
<tr>
<td><strong>Unaudited</strong></td>
<td>Indicates the item has explicitly been marked as not audited. Items processed in Nuix 3.0 will not have this flag set for unaudited items.</td>
</tr>
<tr>
<td><strong>URI</strong></td>
<td>The URI of the data item. Loosely, shows the location which the item was loaded from.&lt;br&gt;getUri()</td>
</tr>
</tbody>
</table>

**Note**: (*) The audited size is generally the size of the item itself for loose files or in the case of email messages, the size of attachments is excluded from the audited size of the email item itself. In the case of containers, the container item itself will have an audited size of 0. (***) Metadata items marked with a double asterisk might be slow to compute, especially over large datasets, and should therefore be avoided on the Results view and during exports. Writing this metadata to custom metadata fields may assist with slowness when using these fields in profiles.