# Remote Matching Service

## User's Guide

Version: 2.1.x

Written by: Documentation Team, R&D
Date: Thursday, June 24, 2021

Hyland

# Copyright

# Table of Contents

# Overview

The Remote Matching Service (RMS) provides centralized search services to applications such as Brainware Intelligent Capture (BIC) Runtime Server and BIC Verifier, and a web-based administration client. RMS is also capable of supporting large search data pools.

# The Administration Client

The web-based Administration Client allows you to monitor and configure the system.

## Log in

You can connect to the Administration Client on any server where RMS is installed.

Typically, the client URL is `http://[servername]:8080/RMS/admin/`.

**Note:**
This URL is case-sensitive.

Unless your installation is protected by a Single Sign-On (SSO) provider, you are redirected to a login screen when trying to access the Administration Client.

When the Remote Matching Service is protected by an SSO provider, the SSO system will handle the authentication.

**Note:** When your installation of RMS is connected to an installation of Brainware Intelligent Capture, you must be in the administrator role in order to be authorized to use the Administration Client.

To log into the Administration Client, complete the following steps.

1. In the login screen, in the **User name** field, type your user name.
2. In the **Password** field, type your password.
3. Click **Connect**.

   If your credentials are accepted and you are authorized to use the application, a dashboard appears. If you have entered invalid credentials or if you are not authorized to use the application, an error message appears.

## Terminate a session

Your session terminates automatically after a period of inactivity.

To manually terminate a session, do the following.

- In the upper right corner of the window, click  and then select 

# Navigation

## Navigation pane

You can use the navigation pane to switch between options and to select items to work on, such as data sources, search indexes and server instances.

You can show or hide the navigation pane to free up space for the selected option. This can be useful when viewing a duplicates report, or when searching.

To hide or show the navigation pane, click the  icon on the toolbar.

## Action bar

The action bar contains the  button to hide or show the navigation pane, and the buttons that are available for the selected data source, search index or server instance

## Context menus

Right-click on a data source, search index or server instance to display a menu of options that are relevant to the item. There is no context menu for the Dashboard, Messages, Search, Duplicates or any of the Settings options. The options in the context menu are identical to those in the action bar.

# About the dashboard

The Dashboard displays an overview of the RMS system highlighting potential problems that need investigation. The state of the system refreshes automatically. The left pane shows the state of the configured servers and the right pane shows the state of the search indexes.

To show the available items with which you can work, expand by clicking  and collapse by clicking  next to the **Data Sources**, **Search Indexes**, **Server Instances** and **Settings** options. The last selected item or the first available item are shown.

## Check server state

In the left pane of the Dashboard, an entry for each configured server displays, except for the load balancer. The host name and port of each server shows with an icon indicating the server state.

The  icon indicates the server is available.

The  icon indicates the server is unavailable. It could be offline or incorrectly configured.

The  icon displays while the RMS is determining the state.

To display server details for a server instance, complete the following step.

- On the **Dashboard**, in the **Servers** pane, click the name of the server that you want to view.

## Check index state

In the right pane of the Dashboard, an entry for each configured search index displays. The name, index state, index status and last rebuild start time is shown for each index along with an icon that indicates the overall state of the index.

The  icon indicates a status of good, meaning an index is open, and a rebuild succeeded or is in progress.

The  icon indicates an index was manually closed.

The  icon indicates an index was unloaded after a period of inactivity.

The  icon indicates an index failed to rebuild or the index state is not open on any server.

To display index details, complete the following step.

- On the **Dashboard**, in the **Indexes** pane, click the name of the index that you want to view.

## Index states

Additional information in the Dashboard shows the state of the index on each server. An Index State can be one of the following.

- Open. The index is available for search requests.
- Opening. The index will shortly be available for search requests.
- Unavailable. The index is not available. The server instance may be down.
- Maintenance. The index exists and is available on the server, but cannot be used for searching currently. This may be the case if the index is being recreated or if RMS is about to bring the index up-to-date after the server instance is started.
- Closing. The index is in the process of closing down.
- Closed. The index is closed and unavailable for searching.
- Closed (inactive). The index is closed following a period of inactivity. Searching on the index will automatically cause the index to open.

## Indexer status

Additional information in the Dashboard shows the status of the component that is responsible for rebuilds. An Indexer Status can be one of the following.

- Succeeded. The last rebuild completed successfully.
- Failed. The last rebuild failed.
- In progress. A rebuild is currently in progress.
- Ready for commit. The rebuild has successfully completed on a machine, but other machines are still processing.
- Not started. The index has not been built since it was created.
- Unavailable. The indexer status is not available. The server is either offline or not configured correctly.

- Imported: The index is imported from another system.

## About configuring data sources

You must configure a data source to import records from a database to your RMS system.

A data source can be used with any server instance. You can set up a data source using a process similar to setting up an ODBC data source in Windows.

By default, RMS is compatible with Microsoft SQL Server and Oracle databases. RMS can connect to other databases if a JDBC driver is available for that type of database.

When you configure an RMS data source, you must specify a name for the connection. Once you submit a name for a data source in RMS, you cannot modify the name because search indexes reference the data source under the specified name.

You can only delete a data source if the data source is not currently referenced by any search index or used to configure reporting.

### Configure a Microsoft SQL Server data source

To configure a data source based on a Microsoft SQL Server (MSSQL) database, complete the following steps.

**Prerequisite**

To use a MSSQL Server data source, you must have Microsoft SQL Server authentication enabled and TCP/IP enabled on your system. The value is usually 1433 for SQL Server and 1521 for Oracle.

1. On the **navigation** pane, expand **Data Sources**.
2. Right-click a data source and select **Add > SQL Server**.
3. In the **Add Data Source** dialog box, in the **Configuration** section, complete the following substeps.
    1. In the **Name** field, type a name for the data source.
    2. In the **Host** field, type the server host name or IP address.
    3. In the **Port** field, type the TCP/IP port number. The value is typically `1433`, which is the default setting.
    4. (Optional) In the **Database** field, type the name for the default database for RMS to use to connect to the data source. If you leave this field blank, the default database is the database assigned to the user account that connects to the database.
4. (Optional) To test the data source configuration, in the **Add Data Source** dialog box, in the **Test Configuration** section, complete the following substeps.
    1. In the **User Name** field, type the database user name.
    2. In the **Password** field, type the database user password.
    3. Click **Test**.

        RMS tests the connection and displays a success or failure message.
5. Click **Save** to create the data source with the current configuration.

## Configure an Oracle data source

To configure a data source based on an Oracle database, complete the following steps.

1. On the **Navigation** pane, expand **Data Sources**.
2. Right-click a data source and select **Add > Oracle**.
3. In the **Add Data Source** dialog box, in the **Configuration** section, complete the following substeps.
    1. In the **Name** field, type a name for the data source.
    2. In the **Host** field, type the server host name or IP address.
    3. In the **Port** field, type the TCP/IP port number for which the database is available. The value is typically `1521`, which is the default setting.
    4. In the **SID** field, type the Oracle System ID for the data source.
4. (Optional) To test the data source configuration, in the **Add Data Source** dialog box, in the **Test Configuration** section, complete the following substeps.
    1. In the **User Name** field, type the database user name.
    2. In the **Password** field, type the database user password.
    3. Click **Test**.

        RMS tests the connection and displays a success or failure message.
5. Click **Save** to create the data source with the current configuration.

## Configure a generic data source

To configure a data source based on a database from a provider other than Microsoft SQL Server or Oracle, complete the following steps.

**Prerequisite**

To use such a database, you must obtain a JDBC driver for that database.

1. Copy the JAR file containing the JDBC driver to the *WEB-INF\lib* directory used by the RMS web application on all server instances.
2. On the **Navigation** pane, expand **Data Sources**.
3. Right-click a data source and select **Add > Generic**.
4. In the **Add Data Source** dialog box, in the **Configuration** section, complete the following substeps.
    1. In the **Name** field, type a name for the data source.
    2. In the **Driver Class** field, type the class name of the JDBC driver.

        **Example**

        For example, `org.apache.derby.jdbc.EmbeddedDriver` for an Apache Derby database
    3. In the **JDBC URL** field, type the URL for the JDBC driver for the data source.

        **Example**

        For example, `jdbc:derby:databaseName` for an Apache Derby database
5. (Optional) To test the data source configuration, in the **Add Data Source** dialog box, in the **Test Configuration** section, complete the following substeps.

1. In the **User name** field, type the database user name.
2. In the **Password** field, type the database user password.
3. Click **Test**.

   RMS tests the connection and displays a success or failure message.

6. Click **Save** to create the data source with the current configuration.

### Modify a data source

To modify a data source, complete the following steps.

1. On the **navigation** pane, expand **Data Sources**.
2. Right-click the name of the data source you want to modify and select **Modify**.
3. In the **Modify Data Source Details** dialog box, edit any of the configurable fields.

   **Note:** The data source name cannot be changed.

4. (Optional) To test the modified data source configuration, in the **Modify Data Source Details** dialog box, complete the following substeps.
   1. In the **User name** field, type the database user name.
   2. In the **Password** field, type the database user password.
   3. Click **Test**.

      RMS tests the connection and displays a success or failure message.
5. Click **Save** to apply your changes.

### Delete a data source

To delete a data source from the system, complete the following steps.

1. On the **navigation** pane, expand **Data Sources**.
2. Right-click the name of the data source you want to delete, and select **Delete**.
3. In the **Delete Data Source** dialog box, click **Yes**.

   **Note:** If you attempt to delete a data source that is currently used by a search index or to configure reporting, an error message appears.

## About search indexes

RMS organizes collections of searchable data records, such as vendor addresses, in search indexes that can be accessed independently of each other. Clients like Brainware Intelligent Capture connect to a single search index that contains the data for the field you are analyzing.

A search index is connected to a record source – either a CSV file or a database – and builds itself periodically based on the records retrieved from that source.

Optionally, a CSV file can be provided that contains records that will be excluded from the search. If an ID column is specified for the search index, records that need to be excluded are identified by the ID. Otherwise, the records in the excluded records CSV file need to be identical to the ones in the main record source.

Search indexes consume system resources while they are loaded. Optionally, search indexes can be unloaded until they are required; this can be done either manually or after a period of inactivity. When an index has been unloaded due to inactivity, searching on the index will cause its resources to be reloaded – the search will not fail but the first search may take longer than usual as the index is loaded.

Search indexes can be exported for use on other systems. This allows the search index to be distributed without giving full access to the original records. When these pre-built search indexes are imported, only the scoring mechanism can be configured. For security reasons, imported search indexes cannot be saved or exported, and they cannot be searched for duplicates.

Search index files can be stored with Advanced Encryption Standard (AES) encryption for greater security. A 256 bit key is used if supported, otherwise a 128 bit key is used. The Java Cryptography Extension (JCE) must be installed in the Java Runtime Environment to use 256 bit keys. The JCE can be downloaded from Oracle.

## Scoring configuration

Your site may have scoring configuration enabled. Typically, if client connections request candidates from RMS to compute a final score, scoring is not enabled. Two scoring models are available.

- ASSA-based scoring converts candidate records into a single text and tries to find that in the examined document. You can configure the weight of the correct word order within the score.
- Advanced record scoring allows you to define weights for individual columns in order to give important columns (like a company name) a greater influence on the score than less important ones (like a telephone number). You can also treat two or more columns as one if you expect to find their values close to each other in the examined documents (for example ZIP code and city).

## Add a search index

To add a new search index, complete the following steps.

1. On the **navigation** pane, expand **Search Indexes**.
2. Right-click and click **Add**.
3. Do any of the following.
   - Select **CSV File** if your data is saved in a text file that contains a list of data values separated by a particular character.
   - Select **Database**, if your data is located in a database.

   For information on specific fields, refer to the table below. After you have provided all mandatory information, RMS tries to read the first rows of the file based on the current configuration (for CSV File) or tries to select the first rows from the database based on the current configuration (for Database) and displays a confirmation message.

| Field | Description |
|---|---|
| Index Name | Type a unique name for the index.<br>**Note:**<br>This field is applicable for both **CSV File** and **Database** options. |
| **Include**<br>**Note:**<br>The **Source File**, **first row contains column names**, **Separator**, and **Quote Character** fields are applicable for the **CSV File** option only.<br>The **Data Source**, **Username**, **Password**, and **Query** fields are applicable for the **Database** option only. | |
| Source File | Type the path of your CSV file location.<br>**Note:**<br>When working with multiple server instances, make sure to use a path that is accessible by all servers. |
| first row contains column names | Select this check box if the first row of the CSV file contains column names. |
| Separator | Select the character that separates fields in the CSV file. This is typically a semicolon or a comma. |
| Quote Character | Select the character that is used to quote the content of a field. If your CSV file does not have the content of fields in quotes, select **None**. |
| Data Source | Select the data source that contains the data you want to add to the index. The list contains all data sources that have been configured in **Data Sources**. |

| Field | Description |
|---|---|
| User Name | Type the user name that is used to access the data source. |
| Password | Type the password that is used to access the data source. |
| Query | Type in the SQL query that is used to select the records you want to add to the index. For example: `select * from vendor` |
| **Exclude**<br><br>**Note:**<br>The fields in the **Exclude** area of the **Add Search Index** dialog box are applicable to both the **CSV File** and the **Database** options. | |
| Source File | Type in the path to your CSV file of records to exclude.<br><br>**Note:**<br>When working with multiple server instances, make sure to use a path that is accessible by all servers. |
| first row contains column names | Select this check box if the first row of the CSV file contains column names. |
| Separator | Select the character that separates fields in the CSV file. This is typically a semicolon or a comma. |
| Quote Character | Select the character that is used to quote the content of a field. If your CSV file does not have the content of fields in quotes, select **None**. |

4. Select the **Columns** tab, and verify or select the following information.

   This tab is only enabled if you have entered a valid record source configuration.

   - Verify that the displayed column names reflect the available data in your record source by checking that all of the columns are referenced in the SELECT statement or, for CSV files, that the separator and quote characters reflect the contents of the file.

- Select the **Search** check box for all columns whose values should be part of the index. This should include any information you are expecting to find in the analyzed text.
- Select the **Filter** check box for all columns that can be used as a filter to preselect records during a search.
- Select the **ID** button for the column that contains the unique id of the record. This is mandatory when working with Brainware Intelligent Capture. Custom client applications may not need this.
- Select the **Vendor Type** button for the column that contains the vendor type information. This is only required when working with Brainware Intelligent Capture. Refer to the *Brainware Intelligent Capture* documentation for more information about vendor types.

5. Select the **Search Index** tab, and enter the following information.
   - If the ASSA engine is available at your installation, in the **Engine** list, select the search engine you want to use. Record Search is the best choice in most cases. Choose ASSA when migrating from an earlier version of Brainware Intelligent Capture or when working with Chinese, Japanese or Korean (CJK) characters.
   - When using the ASSA engine, in the **Engine Instances** field, select the number of instances that will be opened for searching. While a higher number of instances allow more concurrent searches, this value should always be less than or equal to the number of CPU cores you have available on your servers.
   - When using Record Search, check **Merge Digit Blocks** if spaces in consecutive blocks of numbers should be ignored. For example, 123 456 789 would be transformed to 123456789 in the index.
   - In the **Rebuild Time** field, select a date and a start time for the rebuild of the index. Use the date format year-month-day, such as 2019-11-01. To schedule automatic rebuilds to take place periodically, click **Repeat** and type a time interval. For example, to start a rebuild every night at 10 pm., enter a start date and 22:00 as the time and select 1 Days as the period. The rebuild time is configured in your browser's current time zone not the server's time zone. Internally, the rebuild time is stored in UTC (Coordinated Universal Time).
   - If you want the index to be rebuilt only if it changes when the scheduled time arrives, click **Rebuild Only if Changed**.

     For CSV record sources, the rebuild will only occur if the original file used to create the search index has been modified since the last rebuild.

     For database record sources, either select **Stored Procedure** from the list and enter the name of a stored procedure to run, or select **File Modified** from the list and enter the location of a file whose modification time should be checked. The stored procedure needs to check if something has been modified since a given time. It must accept two parameters. The first is an input parameter with the time to check; i.e., the time the index was last built. The second parameter is an output parameter that should be set to 0 if nothing has changed or a non-zero value if something has changed. For example:

     ```
     create procedure VendorChanged(@ts datetime, @cnt int out) as begin
     select @cnt = count(*) as cnt from vendor_update where last_update
     > @ts end
     ```

   - To load the search index when the server starts up, select the **Load** check box. If you want the

index to be unloaded automatically after a period of inactivity, check the **Automatically unload when inactive for** check box and select the period of time after which the index should be unloaded.

**Note:**
The Dashboard may not reflect the closed state immediately.

- Select the **Encrypt Index Files** check box if the index files are to be stored in an encrypted format.

**Note:**
The index will not reflect this setting until it has been rebuilt.

6. On the **Scoring** tab, enter the following information.

**Note:** This tab is only visible if scoring is available at your installation.

**Example**

If ASSA scoring is available:

1. In the **Scoring Type** list, select the scoring type that you want to use.
   For cases where you want fine control over the scoring mechanism, select **Advanced Record**.

**Example**

To configure ASSA-based scoring, complete the following steps.

1. In the **Word Order Weight** field, type an integer number between 0 and 100 that reflects the influence that the order of search words in the record has on the resulting score. The higher the word order weight, the more important the word order is.
2. Select the **Merge Digit Blocks** check box if spaces in consecutive blocks of numbers should be ignored. For example, 123 456 789 would be transformed to 123456789 in the index.

**Example**

To configure Advanced Record scoring, complete the following steps.

**Note:**
You need to use Advanced Record scoring if ASSA-based scoring is not available.

1. In the **Weighting Mode** list, select the weighting mode that you want to use.
   In **Relative mode**, the weight of a column within the score is basically determined by the length of the value – a long name has a higher influence than a short ZIP code. You can increase or decrease the weight of individual columns relative to other columns. For example, to give a customer number a higher weight you can assign a higher value. The default weight for each column is one.

In **Absolute mode**, column weights are defined as absolute percentages. You may, for example, give a customer number a weight of 30% - if that number is not found the score cannot be higher than 70% even if everything else matches.

2. For Absolute weighting mode, either select the **Distribute weight if value is missing** check box to distribute the weight of a field with no value to the other fields, or clear it and in the **Missing Value Score** field, type an integer between 0 and 100 that indicates the weight given to fields that have no value. This allows the score to be reduced for records that have several missing values. When the value is zero, a missing value adds nothing to the score and the value is considered as not being found at all. When it is 100; it adds the full weight of the column to the score and the value is considered as found.
3. Select **Merge Digit Blocks** if spaces in consecutive blocks of numbers should be ignored. For example, 123 456 789 would be transformed to 123456789 in the index.

**Example**

The column scoring grid has a row for each column. To edit a value click on the relevant cell. For information on specific values, refer to the table below.

| Value | Description |
|---|---|
| **Weight** | Enter a percentage value between 0 and 100 when using Absolute weighting. The total weight across all columns should add up to 100%. If the total weight is more than 100%, your search results will be normalized onto the range 0 – 100%. For Relative weighting mode, the value is relative to the values in the other columns. |
| **Min. Score** | Enter a threshold value for column scores. If the score for a value is below that threshold, the value is considered as not found at all. For instance, on a numeric customer number, you could set a minimum score of 90% but on a customer name field, a value of 60% would be more tolerant of errors in the name. The aim is to avoid, for example, a numeric field with some numbers in common with a numeric search term from increasing the score when it is obviously a false match. |
| **Optional** | This field is available for Relative weighting only. Select the check box if the field score should only go into the total record score if its |

| Value | Description |
|---|---|
| | score is beyond the Min. Score column's threshold value, otherwise its value is ignored. For example, for a telephone number field, it's significant if the value is present but doesn't reflect negatively if it's not. |
| **Word Order** | Select the check box if the order of the words is important and should be checked. For example, word order may not be significant for names – John Smith and Smith, John are equivalent – but street names might need to be ignored completely if the word order is wrong. |

**Example**

To merge the values of other columns with a column in order to treat them as one value, complete the following steps.

- To select the columns to merge, click **Merge cell** and select the **Available Columns** whose text is to be merged with this one.
- To change the order of the columns in the merged text, select one of the **Merged Columns** and use the up and down buttons to change the order of the text.
- To save the configuration, click the  button to save the configuration.

**Note:**
If none of the **Available Columns** are selected, this column is not merged with any others. A searchable column can be merged with more than one column, however you cannot merge a column with itself or with columns that are already merged.

7. Click **Save**.

   When configuring a new index, you will be asked if you want to rebuild the index immediately. If you reject this, the index will not be available for searches until the first scheduled rebuild has succeeded.

### Search indexes and a filter column

When working with a Record Search index and a filter column, try to sort the records in the CSV file or in the SQL query by the column that is used as a filter. This helps to improve the performance especially when working with a larger number of records.

### ASSA search indexes with large amounts of data

When working with an ASSA index and large amounts of data, ASSA will distribute the data to multiple smaller ASSA indexes that contain no more than a million records each; these indexes will be searched in parallel.

While this improves the performance, it also means a higher CPU usage for a single request. Keep this in mind when selecting the number of engine instances. For example, if you are indexing 1.5 million records each request will use two CPU cores instead of one.

If you have just eight cores available, you should not use more than four engine instances because more parallel searches could be active than there are CPU cores available.

## Import a search index

To import a search index, complete the following steps.

1. On the **navigation** pane, expand **Search Indexes**.
2. Right-click **Search Indexes**, click **Add**, and then click **Import**.
3. In the **Source Directory** field, type the location of the directory that contains the exported search index.

   The **Index Name** field will be populated with the exported search indexes that are found in the source directory.
4. In **Index Name** field, select the search index to import.
5. Click **Import Index**.

> **Note:** You will be warned if you are about to overwrite an existing search index or are about to re-import an existing search index.

## Modify a search index

You should consider these items when modifying a search index.

- When you change the column configuration, changes will not be active for clients until the next rebuild is successful.
- For search indexes that are imported, you can only modify settings on the Scoring tab, and the number of engine instances on the Search Index tab for an ASSA search index.

To modify the configuration of a search index, complete the following steps.

1. On the Navigation pane, expand **Search Indexes**.
2. Right-click the name of the data source you want to modify and select **Modify**.

3. In the **Modify Search Index** dialog box, edit any of the configurable fields as needed.
4. Click **Save**.

## Delete a search index

To delete a search index, complete the following steps.

1. On the **Navigation** pane, expand **Search Indexes**.
2. Right-click the name of the search index you want to delete, and select **Delete**.
3. In the **Delete Search Index** dialog box, click **Yes**.

## View search index summary

To display a summary of the most relevant information about a search index, complete the following steps.

1. On the **navigation** pane, click the name of the search index you want to view.
2. In the right pane, view the information about the index.
   - **Record Source**. Information about the record source, such as its type and location.
   - **Index Settings**. Information about the index configuration, such as the engine type and the scheduled rebuild time.
   - **Status**. Provides the indexer status of the component that is responsible for rebuilds. Refer to the Indexer status and Index states sections for possible values.

## Rebuild a search index

When you are using multiple server instances, a rebuild process is started on all instances that are currently online and the new version of the data is not activated until the index rebuild process is complete.

If the rebuild fails on one or more server instances, the new version will be discarded on all server instances so that all server instances are able to work with the same version. If a server is offline for some time and comes back online, the latest version of all search indexes will be copied from another server to bring it up to date.

A search index rebuild process can start automatically in the background according to the scheduling rules configured in the search index configuration. If you need to start a rebuild process immediately, you can start the process manually.

To start a rebuild process manually, complete the following steps.

**Prerequisite**

The previous rebuild process must have finished. A rebuild process cannot be started while another rebuild process is running. Imported search indexes cannot be rebuilt.

1. On the **navigation** pane, expand **Search Indexes**, and click the name of the search index you want to rebuild.
2. In the right pane, in the **Status** section, click **Rebuild**.
3. Optional. Select **Delete index prior to rebuild**, if you want to recreate the index from scratch.

   **Note:** This option should only be used for troubleshooting if the current index is corrupt. With this

option, the index will not be available for search during the rebuild.

## View search index details and statistics

You can view information and statistics for search indexes on a per-server basis. To view index details, complete the following steps.

1. On the **navigation** pane, select **Search Indexes**, and then click the name of the search index for which you want to view details.
2. In the right pane, in the **Status** section, double-click on the row for the server for which you want to view the details.

### Search index details

The Details tab of the Search Index Details dialog box provides information about the index itself and the last rebuild of the index on that server, including the following information.

- The number of indexed records.
- Start date, end date and status of the last rebuild on this machine. If an error occurred during the rebuild on this machine, an error message also displays.

### Statistics index details

The Statistics tab of the Search Index Details dialog box provides performance statistics for the index on the machine, including the following details.

- The number of searches that occurred in the selected period of time.
- Response time. The average time it took to process a search request on the machine within the selected period of time.
- Time in waiting queue. The average time it took to allocate an engine instance to process the search request. This is only available for indexes that use the ASSA engine. If requests are spending most of their time in the waiting queue, consider increasing the number of engine instances for that index or adding additional servers.

### Set display time

By default, statistics are displayed for the last hour. To change the period of time, complete the following steps.

1. From the time unit list, select a unit of **Minutes**, **Hours**, **Days** or **Weeks**.
2. In the text field, type the required value.
3. Click **Refresh**.

## Save search index data to a file

The accumulated search data of a search index can be saved to a CSV file to use when troubleshooting.

**Note:** You cannot save an imported search index. You do not want to perform this operation if you or another user are currently performing a save operation for the selected search index.

To save data from a search index to a file, complete the following steps.

1. On the **navigation** pane, select **Search Indexes**.
2. Click the name of the search index you want to save.
3. In the **Status** section of the **Search Index Detail** page, double-click the name of the server from which the data is to be saved.
4. Click **Save Index**.
5. In the **Save Index Data** dialog box, complete the following substeps.
    1. In the **Target Directory** field, type the target folder location for the CSV file.
    2. In the **Maximum File Size (MB)** field, type the maximum size for each individual file created during the save.

        **Note:** When one file reaches the specified file size limit, the system creates subsequent files until the remaining data is saved.
6. Click **Save**.

## Export search index data to a file

You can export search index data to a file on a server in your network and securely distribute it to other people.

**Note:** You cannot export an imported search index.

To export data from a search index to a file, complete the following steps.

1. On the **navigation** pane, select **Search Indexes**.
2. Click the name of the search index you want to save.
3. In the **Status** section of the **Search Index Detail** page, double-click the name of the server from which the data is to be exported.
4. Click **Export Index**.
5. In the **Export Index** dialog box, in the **Target directory** field, type the target folder location for the export file.
6. Click **Export Index**.

# About managing server instances

Maintaining multiple server instances allows you to achieve ideal load balancing and to have a failsafe server as a backup if any server becomes unavailable. You can use multiple server instances with RMS. The system automatically detects most server instances every time it starts up.

All active server instances that are connected to the same configuration database provide exactly the same functionality. However, using multiple configured server instances results in better throughput with searches powered by RMS.

If you have millions of addresses or documents in a search pool, the system utilizes the available server instances to achieve faster searches. You must register any separate load balancers with your RMS system to allow the benefit of having multiple server instances. The load balancer can be either another server instance configured to be the load balancer or a third party product.

## Add a new server instance

Server instances are typically added during the RMS installation. After installation, to add a server instance for RMS to utilize, complete the following steps.

1. On the **navigation** pane, select **Server Instances**.
2. Right-click a server instance name, and select **Add**.
3. In the **Add Server Instance** dialog box, complete the following substeps.
    1. In the **Configuration** section, in the **Protocol** list, select either HTTP or HTTPS.
    2. In the **Host** field, type the host name or IP address for your server instance. For example, type `rms_1`.
    3. In the **Port** field, type the port number for your server instance. For example, type `6000`.
    4. In the **Path** field, type the remainder of the path for the server instance. For example, type `/rms/service`.

        **Note:** The path should begin with a slash. RMS adds a slash automatically if one is omitted.

    5. Optional. If the server instance you are adding will function as the load balancer for your system, select the **Load Balancer** check box.

        **Note:** If you have already designated a load balancer for the system, this check box is disabled. To designate a different load balancer, you must delete the old load balancer from RMS.

4. Click **Save**.

## Modify a server instance

To modify the configuration of a server instance, complete the following steps.

1. On the **navigation** pane, select **Server Instances**.
2. Right-click a server instance name, and select **Modify**.
3. In the **Modify Server Instances** dialog box, complete the following substeps.
    1. In the **Configuration** section, in the **Protocol list**, select either HTTP or HTTPS.

2. In the **Host** field, type the host name or IP address for your server instance. For example, type `rms_1`.
3. In the **Port** field, type the port number for your server instance. For example, type `6000`.
4. In the **Path** field, type the remainder of the path for the server instance. For example, type `/rms/service`.

   **Note:** The path should begin with a slash. RMS adds a slash automatically if one is omitted.

5. Optional. If the server instance you are modifying will function as the load balancer for the system, select the **Load Balancer** check box.

   **Note:** If you have already designated a load balancer for the system, this check box is disabled. To designate a different load balancer, you must delete the old load balancer from RMS.

4. Click **Save**.

### Delete a server instance

Deleting a server instance from the system removes the server registration from the configuration database. This process does not uninstall the server from your environment.

To delete a server instance from the system, complete the following steps.

**Prerequisite**

The server instance must be offline prior to deletion.

1. On the **navigation** pane, select **Server Instances**.
2. Right-click the name of the server instance you want to delete, and select **Delete**.

   **Note:** There may be a brief delay as the system verifies the server instance is offline.

## About viewing system messages

RMS stores information about the state of the system. This can include the success or failure of a search index rebuild, or the availability of a server instance. You can view, filter and sort messages to help determine the reason for errors.

The messages are at one of three levels.

- Error. A serious condition was detected which could compromise the running of RMS. For example, a server instance was unavailable.
- Warning. A situation was detected that could affect the performance of RMS.
- Information. Provides details of normal system operation. For example, a server instance started, or a search index rebuilt successfully.

The message includes the server name experiencing the issue.

### View system messages

To view system messages, complete the following steps.

1. On the **navigation** pane, click **Messages**, and select one of these options.
   - In the **Level** list, select the minimum message level you want to show. Error shows error messages, Warning shows both warning and error messages, and Information shows all messages types.
   - In the **Host** list, select the server instance that you want to show. A blank entry shows messages from all server instances.
2. In the **Creation Date** field, type the date or date and time of the earliest message you want to show, and then click **Search**.
3. The system returns the number of messages that match the search criteria. The resulting messages are grouped into pages and the number of pages is shown. To navigate through the pages, complete the following substeps.
   1. Click  to move to the first page of messages.
   2. Click  to move to the previous page of messages.
   3. Click  to move to the last page of messages.
   4. Click  to move to the next page of messages.

      **Note:** Buttons are only enabled when they are relevant. For example, the previous page button is disabled if you are at the first page.

## About searching

The search feature of RMS allows you to query individual servers and indexes, and to retrieve individual records by their IDs, however it is not intended to be used as an end-user search client.

### Perform a search

To perform a search, complete the following steps.

1. On the **navigation** pane, select **Search**, and then click the **Search by Text** tab.
2. In the **Server** list, select the server to run the query on.
3. In the **Index** list, select the index to run the query on. If the index has filters defined, these are displayed in a list. The results grid headings change to reflect the columns in the selected index. If the index has been unloaded due to inactivity, initiating a search will load the index; the first search may take longer while the load occurs.
4. In the **Search by Text** field, type or paste the query text.
5. In the **Search for Records** list, select the type of search.
   - **Containing the Text**. This is appropriate for short queries where the entire text is expected to be within a record.

- **Appearing Within the Text (All Candidates)**. This is appropriate for longer queries where one or more of the indexed records are expected to be found in the query text. Only search indexes that use the Record Search engine support optimized scoring for long queries that contain records.
- **Appearing Within the Text (Best Candidates Only)**. This is also appropriate for longer queries. While the All Candidates option just does a quick but not very precise pre-selection of candidates, an additional verification step boosts the scores and eliminates candidates that are unlikely to match.

6. Optional. If the selected index has filters, select a filter from the **Filter by ID** list and type a value in the text field.
7. Optional. In the **Results** field, type the maximum number of results you require.
8. Click **Search**. The results grid displays the records that match your query.
9. Optional. Click **Reset** to clear the text field and set everything back to its default value.

### Copy results to clipboard

It is often useful to select the concatenated text of a search result or a particular column and copy it to the clipboard for use in RMS or another application.

To copy a row or column to the clipboard, perform a search, and then complete the following step.

- Right click on a search result and click either Copy Row to Clipboard to copy a row, or click Copy Cell to Clipboard to copy a cell value.

### Analyze the score

An analysis of the scoring is available for search results that were generated by Records Appearing Within the Text (Best Candidates Only) for indexes with Advanced Scoring. This shows how the weighting of the columns affected the final score, which is useful when optimizing the scoring configuration. For each column that has a weight defined in the Scoring tab, the score and weight are displayed along with the text from the original record, and the best match that was found.

To view how a search result was scored, perform a search with Appearing Within the Text (Best Candidates Only) selected in the Search for Records field, then complete the following step.

- Either select a search result and click the Scoring button, or right click on a search result and then click Show Scoring Details.

## Retrieve a record

To retrieve a record, complete the following steps.

1. On the **navigation** pane, select **Search**, and then click the **Search by ID** tab.
2. In the **Server** list, select the server to run the query on.
3. In the **Index** list, select the index to run the query on.
4. In the text field, type the ID of the record.
5. Click **Search**.
6. Optional. Click **Reset** to clear the text field and set everything back to its default value.

# About duplicate detection

RMS can create a report that identifies potential duplicate records in the search index repository. The information in the report can be used to remove unwanted duplicate records resulting in reduced space requirements and improved search performance. RMS reports potential duplicate records, but will not delete any records.

Duplicate detection is far more efficient with Record Search indexes. If you have an ASSA search index with more than a hundred thousand records, it is advisable to create a temporary Record Search index from the same record source and to detect duplicates on that before removing the duplicates.

Duplicates cannot be identified in imported search indexes or in search indexes that are not loaded.

## About duplicates reports

A duplicates report consists of two types of rows: an original (which has no entry in the Score column), and a duplicate which has a percentage Score indicating the duplicate row's similarity to the original. An original row may have one or many duplicate rows.

## Create a duplicates report

To detect potential duplicates and create a report, complete the following steps.

1. On the **navigation** pane, select **Duplicates**.
2. In the **Server** list, select the server to run the duplication detection on.
3. In the **Index** list, select the index to run the duplication detection on. Imported search indexes do not appear in the list nor do search indexes that are not loaded.
4. Click the **Create** button, and in the **Create Report** dialog box, complete the following substeps.
    1. In the **Minimum Score %** box, type or select the value that the record must score to be considered a duplicate.
    2. Select the **Boost Queries** check box to allow all of the CPU power to be used for duplicate detection.
    3. For each searchable column, type a relative weight to determine the importance of the column in the scoring process.
5. Click **Create**.

    The report generation begins and the status bar shows the progress of the duplicate detection. This operation may take some time depending on the number of records in your index.

    **Note:** You can work with other RMS features while duplicate detection is in progress.

## Stop duplicate report creation

To stop the generation of the duplicate report, complete the following steps.

1. On the **navigation** pane, select **Duplicates**.
2. In the **Server** list, select the server that ran duplication detection.
3. In the **Index** list, select the index against which duplication detection was run.

4. Click **Stop**.
5. In the **Stop Report** dialog box, click **Yes**.

> **Note:** Stopping the operation does not create a partial report.

## View a duplicates report

To view a duplicates report, complete the following steps.

1. On the **navigation** pane, select **Duplicates**.
2. In the **Server** list, select the server that ran duplication detection.
3. In the **Index** list, select the index against which duplication detection was run.
4. The duplicate report displays, if it exists. If the status bar shows that there is no report available, you need to create a report. The number of original records and potential duplicates displays. The results are grouped into pages and the number of pages is shown.
5. To navigate through the pages, complete the following substeps.
    1. Click  to move to the first page of the list of duplicates.
    2. Click  to move to the previous page of the list of duplicates.
    3. Click  to move to the last page of the list of duplicates.
    4. Click  to move to the next page of the list of duplicates.

    > **Note:** Buttons are only enabled when they are relevant. For example, the previous page button is disabled if you are at the first page.

## Delete a duplicates report

To delete a duplicates report, complete the following steps.

1. On the **navigation** pane, select **Duplicates**.
2. In the **Server** list, select the server that ran duplication detection.
3. In the **Index** list, select the index against which duplication detection was run.
4. Click **Delete**.
5. In the **Delete Report** dialog box, click **Yes**.

## Download a duplicates report

To download a duplicates report, complete the following steps.

1. On the **navigation** pane, select **Duplicates**.
2. In the **Server** list, select the server that ran duplication detection.
3. In the **Index** list, select the index against which duplication detection was run.
4. Click **Download**.
5. Select **Download as Excel** to download the duplicates report as an Excel file, or click **Download as**

**CSV** to the download the duplicates report as a comma-separated file. RMS saves the file to your specified location.

## About configuring Remote Matching Service

You can configure various aspects of RMS to work in your environment.

This section includes information on configuring a mail server, email notifications, reports, users, authentication, and connections to use with RMS.

### Configure client connection

To use RMS with an external client such as Brainware Intelligent Capture, on the navigation pane, select **Settings > Client Connection**.

The required settings for the client application appears.

### Configure a mail server

Configuring a mail server allows RMS to send email notifications that can alert you to important changes in the state of the system such as the outcome of a search index rebuild, or an unavailable server. To configure a mail server, complete the following steps.

1. On the navigation pane, select **Settings > Mail Server**.
2. In the **Host** field, type the host name or IP address for your mail server. For example, type `smtp.yourcompany.com`.
3. In the **Port** field, type the port number for your mail server. For example, type `25`.
4. Optional. If the mail server requires additional authentication, in the **User Name** field, type the user name required to authenticate you to the mail server.
5. Optional. If the mail server requires additional authentication, in the **Password** field, type the password required to authenticate you to the mail server.
6. In the **Sender Address** field, type a valid email address from which any emails will be sent.
7. To test the mail server configuration, complete the following sub-steps.
    1. In the **Recipient Address** field, type a valid email address to receive the test message. For example, type `auser@yourcompany.com`.
    2. Click **Test**.

       RMS sends an email to the recipient and displays a confirmation message.
8. Click **OK**.
9. If the test email could not be sent, check that you have filled in each of the fields correctly. If the email still cannot be sent, check the server log files for more detailed information.
10. Click **Save**.

**Note:**

Even if the test email sends successfully, you must confirm that it arrived with the intended recipient. If the email did not arrive, check the Recipient Address field. The email may have been automatically filed in the addressee's spam folder.

## Configure email notifications

To configure which notifications will be sent by email, complete the following steps.

**Prerequisite**

A mail server must be configured prior to configuring email notifications.

1. On the **Navigation** pane, select **Settings > Notifications**.
2. In the **Email Recipients** field, type a valid email address or a list of valid email addresses separated by commas. For example, type `auser@yourcompany.com, anotheruser@yourcompany.com`.
3. Select the check box next to any of these notifications that you wish to send to the email recipients.
   - **Search Index Rebuild Succeeded**. This notification is sent when a search index is rebuilt successfully.
   - **Search Index Rebuild Failed**. This notification is sent when a search index rebuild fails.
   - **Server Unavailable**. This notification is sent when the system detects that a server cannot be contacted.
4. Click **Save**.

## Configure reporting

Reporting can be configured in RMS to store information about searching and indexing. This information can be used by external reporting tools as the basis for reports.

For each search reporting period, the number of successful and failed searches, and the average, minimum and maximum search times are stored. For search indexing, the time the indexing started; the time taken to index; the total time including the commit; and the number of indexed records are stored.

To configure reporting, complete the following steps.

1. On the **navigation** pane, select **Settings > Reporting**.
2. Select the **Enable Reporting** check box to activate reporting, and then set the following options as needed.
   - In the **Entry Expiration** field, type the number of days that an entry will be available for reporting.
   - In the **Search Reporting Period** field, type the number of minutes that search-based usage information is averaged over. For example, type `5` to store the search statistics every five minutes.
   - If the reporting data is stored in an external repository, in the **Data Source Name** list, select the data source where the data should be written. This list contains all data sources that have been configured in Data Sources.
   - If the reporting data is stored in an external repository and it requires credentials, in the **Data Source User** field, type the user name and in the **Data Source Password** field, type the password.

## About configuring internal users

RMS includes the tools necessary to define and manage the user accounts that are authorized to use its services. The user accounts that it manages directly are referred to as "internal" accounts. RMS can also be configured to use an external authentication service, such as LDAP, or a Single Sign-On (SSO) provider.

Other computer applications and other RMS instances, referred to as technical clients, can access its services only by using an internal account. RMS includes some predefined internal accounts that can be edited, but not deleted. When you are not using an external authentication service, you can create additional internal accounts to log on to the administration client.

When not using an external authentication service, additional internal accounts can be created for human users to log in with the administration client.

### Add a new internal user

To add a new internal user, complete the following steps.

1. On the **navigation** pane, select **Settings > Internal Users**.
2. Right-click an internal user, and select **Add**.
3. In the **Name** field, type a unique name for the internal user.
4. In the **Password** field, type a password for the user. The password field cannot be blank.
5. In the **Confirm Password** field, retype the password for the user.
6. Click **Save**.

### Delete an internal user

To delete an internal user, complete the following steps.

1. On the **navigation** pane, select **Settings > Users**.
2. Right-click the name of the internal user you want to delete, and select **Delete**.
3. In the **Delete User** dialog box, click **Yes**.

### Modify an internal user

To modify the configuration of an internal user, complete the following steps.

1. On the **navigation** pane, select **Settings > Users**, and click the name of the internal user you want to modify. The internal user name field cannot be modified.
2. In the **Password** field, type a new password for the user. The password field cannot be blank.
3. In the **Confirm Password** field, retype the password for the user.
4. Click **Save**.

## About configuring authentication

RMS supports the LDAP and Intelligent Capture external authentication mechanisms. Additionally, Single Sign-On (SSO) authentication can be configured.

### Configure LDAP authentication

LDAP servers (including Microsoft Active Directory) can be used to authenticate users. There are two ways to configure the LDAP connector for RMS.

If all users are within one node of the LDAP tree, a simple template-based mechanism can be used to look up users.

If the users are stored within a deeper tree structure, the connector needs to perform a search operation to resolve a user.

1. On the **navigation** pane, select **Settings > Authentication**.
2. In the **Authentication** section, in the **Authentication Type** field, select **LDAP**.
3. If the users are all within the same node within the LDAP tree, complete the following substeps.
    1. In the **Users Are** field, select in a single node, and then type the URL for the LDAP server and a template to locate the user in **User DN Template**. For example, `uid={user}, ou=myOU, o=myOrg` where `{user}` is a placeholder for the user name in the template.
    2. In the **Server URL** field, type the URL for the LDAP server.
4. If the tree structure needs to be searched to find the user, complete the following substeps.
    1. In the **Users Are** list, select in a tree structure, in the **User DN** field, type the full distinguished name of a user to use for searching the LDAP repository, and in the **Password** field, type the user password for that user.
    2. In the **Search Filter** field, type the user search filter, for example `sAMAccountName={user}` where `{user}` is a placeholder for the user name and must be part of the filter.
    3. Optional. In **Search Paths**, specify the start points for the searches. The entire directory is searched if no paths are specified.
5. Click **Save**.

### Configure Intelligent Capture authentication

RMS can be configured to use the Brainware Intelligent Captures user database for authentication. This means users can log in with the same credentials they can use to log in to Brainware Intelligent Capture.

To configure Brainware Intelligent Capture as the external authentication mechanism, complete the following steps.

1. On the navigation pane, select **Settings > Authentication**.
2. In the **Authentication** section, in the **Authentication Type** field, select **Intelligent Capture**.
3. In the**Database Type** field, select either **SQL Server** or **Oracle**.
4. In the **Host** field, type the name or IP address of the machine that holds the database.
5. In the **Port** field, type the port for the database.
6. In the **User Name** field, type the database user name.

7. In the **Password** field, type the database user password.
8. If you selected an Oracle database, in the **Oracle SID** field, type the identifier for your Oracle system.
9. Click **Save**.

## Configure single sign-on (SSO) authentication

If access to RMS is protected by an SSO provider to supply the name of the authenticated user in an HTTP header field, you can configure the name of that header field in order to accept the SSO authentication.

To configure SSO authentication in the RMS system, complete the following steps.

1. On the **navigation** pane, select **Settings > Authentication**.
2. In the **Single Sign-On** section, in the **HTTP Header** field, type the HTTP header that the single sign-on system uses to pass the logged in user name.
3. Click **Save**.

## Configure authorized users

By default, all internal users are authorized to use RMS. If an external authenticator or single sign-on (SSO) is configured, the system needs to know which users are authorized.

**Note:** You should consider these items when configuring authorized users.

If you are integrating with Brainware Intelligent Capture and that systems user database is configured for authentication, all users with the “Administrator” role are authorized to use RMS.

If you are integrating with Brainware Intelligent Capture, but have SSO or LDAP configured for authentication, all users with the “Administrator” role are authorized.

To provide a list of authorized users for RMS to use, completing the following steps.

1. On the **navigation** pane, select **Settings > Authentication**.
2. In the **Authentication** section, in the **Authorization Type** field, select **Internal Users Only**, **Listed Users** or **Intelligent Capture**.
3. For Listed Users authorization, type a list of authorized users separated by commas in the **Authorized Users** field.
4. For Intelligent Capture authorization, type information as needed.
    1. In the **Database Type** field, select either **SQL Server** or **Oracle**.
    2. In the **Host** field, type the name or IP address of the machine that holds the database.
    3. In the **Port** field, type the port for the database.
    4. In the **User Name** field, type the database user.
    5. In the **Password** field, type the database user password.
5. If you selected an Oracle database, in the **Oracle SID** field, type the identifier for your Oracle system.
6. Click **Save**.

### Configure connections for SSL certificates

When you register server instances with the HTTPS instead of the HTTP protocol, the server instances, by default, will check if the SSL certificates are valid when communicating with each other. These checks must be turned off if you are using self-signed certificates.

To configure connections, complete the following steps.

1. On the **navigation** pane, select **Settings > Connections**.
2. Clear the **Check SSL Certificates** check box, if you are using self-signed SSL certificates.
3. Click **Save**.

## View log files

RMS runs as a web application within an Apache Tomcat container. All logging output writes to the log files of Apache Tomcat.

To view the log file for a server, complete the following steps.

1. Navigate to the directory where Tomcat is installed. For example, *[drive:]\Program Files\Apache Software Foundation\Tomcat 8.0*.
2. In the logs directory, locate the file that contains *stdout* and the date you are interested in viewing and open it using a text editor. For example, `tomcat8-stdout.2015-09-25`.