

IBM StoredIQ Insights

*User Guide*



**Note**

Before using this information and the product it supports, read the information in [Notices](#).

This edition applies to Version 7.6.0.20 of product number 5724M86 and to all subsequent releases and modifications until otherwise indicated in new editions.

© **Copyright International Business Machines Corporation 2019.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

---

# Contents

<b>About this publication.....</b>	<b>iv</b>
IBM StoredIQ product library.....	iv
Contacting IBM StoredIQ customer support.....	iv
<b>IBM StoredIQ Insights.....</b>	<b>1</b>
<b>Working with IBM StoredIQ Insights.....</b>	<b>2</b>
Filtering data.....	3
Document tagging.....	6
Visualizing your data.....	11
<b>Filter condition reference and expression language.....</b>	<b>16</b>
Full-text syntax.....	16
Full-text macros.....	19
Indexed annotations.....	20
Regular expressions.....	21
Extended ASCII characters.....	22
Attribute condition reference.....	23
Specification of attribute conditions.....	24
Type of object to match.....	29
Multiple matching values for operators.....	29
At-least expressions.....	30
Parent and child conditions.....	30
No-inheritance conditions.....	31
Sampling.....	32
<b>Notices.....</b>	<b>33</b>
Trademarks.....	34
Terms and conditions for product documentation.....	35
IBM Online Privacy Statement.....	35
<b>Index.....</b>	<b>37</b>

## About this publication

---

*IBM StoredIQ Insights User Guide* describes how to search for data of interest and evaluate the results.

## IBM StoredIQ product library

---

The following documents are available in the IBM® StoredIQ® product library.

- *IBM StoredIQ Overview Guide*
- *IBM StoredIQ Deployment and Configuration Guide*
- *IBM StoredIQ Data Server Administration Guide*
- *IBM StoredIQ Administrator Administration Guide*
- *IBM StoredIQ Data Workbench User Guide*
- *IBM StoredIQ Policy Manager User Guide*
- *IBM StoredIQ Cognitive Data Assessment User Guide*
- *IBM StoredIQ Insights User Guide*
- *IBM StoredIQ Integration Guide*

## Contacting IBM StoredIQ customer support

---

For IBM StoredIQ technical support or to learn about available service options, contact IBM StoredIQ customer support at this phone number:

- 1-866-227-2068

Or, see the Contact IBM web site at <http://www.ibm.com/contact/us/>.

### **IBM Knowledge Center**

The IBM StoredIQ documentation is available in [IBM Knowledge Center](#).

## Contacting IBM

For general inquiries, call 800-IBM-4YOU (800-426-4968). To contact IBM customer service in the United States or Canada, call 1-800-IBM-SERV (1-800-426-7378).

For more information about how to contact IBM, including TTY service, see the Contact IBM website at <http://www.ibm.com/contact/us/>.

# IBM StoredIQ Insights

IBM StoredIQ Insights provides dynamic and interactive filtering for your data with easy access to all metadata and instant plain-text preview of document content for full-text indexed volumes.

Faceted search lets you drill down to refine your search results as needed. In addition, you can apply any valid IBM StoredIQ filter query. Tags let you categorize the data for easier management. Visual representations of search results help you gain further insights into your data. Several chart types let you look at and explore data from different perspectives, thus helping you identify patterns and relationships very quickly.

With IBM StoredIQ Insights, you can search data that is managed and indexed by a data server of the type DataServer - Distributed. In mixed deployments that have classic and distributed data servers, only the content from distributed data servers will be searchable.

The screenshot displays the IBM StoredIQ Insights web application. The top navigation bar shows 'IBM StoredIQ Insights' and 'IBM Docs with filter > Job'. A search bar on the right contains the text 'Search' and 'Save all as Infobot'. Below the navigation bar, there are tabs for 'Result list', 'Basic charts', and 'Analytics charts'. The 'Result list' tab is active, showing a table of search results. The table has columns for 'Object name', 'Size', 'Volume', 'Path', 'Created', and 'Last modified'. There are 2,305 results displayed. The left sidebar contains a 'Facets' section with various filters like 'Tag', 'Category', 'File name extension', 'Object size', 'Created date', 'Modified date', 'Error code', 'Processing state', 'Volume', and 'Path'. The 'Processing state' filter is expanded, showing 'Content - Successful' (2,349), 'Content - Failed' (8), 'Binary - Not performed' (2,350), and 'System - Successful' (2,350).

Object name	Size	Volume	Path	Created	Last modified
IBMPOF_S55HEC33.pdf #6.0 obj United States English (IBM) Site map (Client IBM) - In: HubSpotplace Submittal IBM Log	472.23 KB	ds1_elastic IBMDocs - Volume	PDF Docs S55HEC	14/11/2019, 07:41:02	12/11/2019, 20:06:18 <a href="#">IBM Products: usabec</a>
IBMPOF_S55EPG77.pdf #6.0 obj United States English (IBM) Site map (IBM Submittal) IBM Site navigation IBM Support	457.53 KB	ds1_elastic IBMDocs - Volume	PDF Docs S55EPG	14/11/2019, 07:35:10	12/11/2019, 21:53:40 <a href="#">IBM Products: usaseg</a>
IBMPOF_S55EPG82.pdf #6.0 obj United States English (IBM) Site map (IBM Submittal) IBM Site navigation IBM Support	15.29 KB	ds1_elastic IBMDocs - Volume	PDF Docs S55EPG	14/11/2019, 07:35:39	12/11/2019, 21:53:42 <a href="#">IBM Products: usaseg</a>
IBMPOF_S55EPG268.pdf #6.0 obj United States English (IBM) Site map (IBM Submittal) IBM Site navigation IBM Support	476.68 KB	ds1_elastic IBMDocs - Volume	PDF Docs S55EPG	14/11/2019, 07:35:29	12/11/2019, 21:55:26 <a href="#">IBM Products: usaseg</a>
IBMPOF_S55HEC38.pdf #6.0 obj United States English (IBM) Site map (IBM Submittal) IBM Site navigation IBM Support	457.74 KB	ds1_elastic IBMDocs - Volume	PDF Docs S55HEC	14/11/2019, 07:40:44	12/11/2019, 20:06:22 <a href="#">IBM Products: usabec</a>
IBMPOF_S55EPG84.pdf #6.0 obj United States English (IBM) Site map (IBM Submittal) IBM Site navigation IBM Support	459.49 KB	ds1_elastic IBMDocs - Volume	PDF Docs S55EPG	14/11/2019, 07:35:33	12/11/2019, 21:53:44 <a href="#">IBM Products: usaseg</a>

---

# Working with IBM StoredIQ Insights

Use IBM StoredIQ Insights to quickly discover relevant data through queries that you can refine as you go. Assign tags to data objects to label and categorize them for easier management. Create visual representations of search results by using the provided chart options.

Before you can work with IBM StoredIQ Insights, you must meet several prerequisites:

- Your deployment must include at least one Elasticsearch cluster that is managed by a data server of the type **DataSet - Distributed**.

In an environment with both types of data server (**DataSet - Distributed** and **DataSet - Classic**), you can work in IBM StoredIQ Insights only with the data that is managed by distributed data servers. Therefore, you might see mismatching object counts for mixed infosets.

- At least, a metadata index must exist. However, some queries return meaningful results only for full-text indexed data or aren't even available if no full-text index exists.
- Access to IBM StoredIQ Insights requires the **Insights User** or the **Insights Reviewer** role.

Users with just these roles have direct access to IBM StoredIQ Insights when they log in to IBM StoredIQ. Users with multiple IBM StoredIQ roles can switch applications by selecting **Insights** from the list of available applications. This list can be expanded by clicking the arrow next to the application name in the global navigation bar.

- IBM StoredIQ Insights administration such as creation of tag groups and tags requires the **Admin** role and the **Insights User** role. Users with these roles can access the administration functions by clicking the user icon and selecting **Insights administration**.
- A screen resolution of 1024 x 768 pixels or higher is recommended.

## Key concepts

*Facets* are high-level categories that represent different characteristics of the data in an infoset. They already give you a rough understanding of the data in the selected infoset even before you run any queries. Facets can be attribute based and non-attribute based. Attribute based facets are object category, file name extension, object size, created date, modified date, volume, and path. Non-attribute based facets are indexed annotations and tags. Each facet has a dynamic set of facet filters, which is derived by ranking the values that are found in the infoset. IBM StoredIQ Insights users can narrow down search results by applying multiple facet filters to find particular data objects.

*Tags* are keywords or labels that IBM StoredIQ Insights users can add to content. They can use such tags to group, manage, and find data objects quickly. The tags are organized in *tag groups*. A carefully thought through tag system provides the basis for drawing the most profit from tagging. Tags are also used when you classify documents by using machine learning models.

*Reviewer tags* are a special kind of tags that can be used to limit access to documents in an infoset. Reviewer tags are maintained in a system-managed **Reviewer** tag group. This tag group shows up as a separate facet in the facet list.

## IBM StoredIQ Insights users

These types of users have access to IBM StoredIQ Insights:

### Administrators

Create and maintain tag groups and tags and can perform the same tasks as regular users

### Regular users

Have access to all data objects for searching, previewing, creating charts, managing tag assignments, and assigning reviewers

## Reviewers

Have access to a limited set of data objects for searching, previewing, and creating charts and can change the reviewer assignment for a data object from their set (thus losing access to the data object)

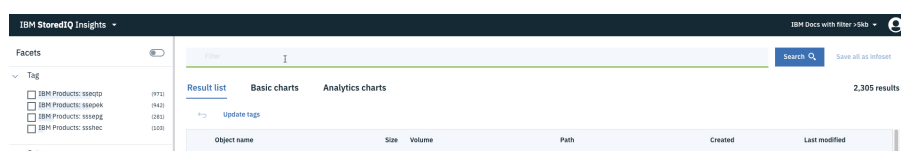
## Filtering data

Select facet filters to filter the data managed by IBM StoredIQ and narrow the search results to compile an info set that represents the data relevant to you.

### What are my search options?

As an IBM StoredIQ Insights user, you can search in multiple ways:

- Select filters from the **Facets** list.
- Enter search terms in the filter field.
- Combine both options.



Facets provide high-level search criteria. They group items and attributes, so called *facet filters*, to target meaningful results. The set of facet filters is dynamic based on the content of the info set. By default, up to six facet filters are displayed. A **Show more** link indicates that further facet filters are available. The maximum number of facet filters that can be displayed varies per facet. The facet filters are usually listed in descending ranking order. Each facet filter shows a *facet count*. The facet count indicates the number or, when you hover over a facet filter, the percentage of data objects in the info set that match the facet filter.

You can choose from the following set of facets:

### Indexed annotation

Select from all indexed annotations that were found in the info set. This facet is available only for info sets on which a cartridge was run.

The maximum number of indexed annotations that can be displayed is 100 although there might be more.

### Tag

Select from all document tags that were found in the info set. This facet is available only if the info set contains tagged data objects.

The maximum number of tags that can be displayed is 50 although there might be more.

### Category

Select from the object categories in the info set.

The maximum number of available facet filters is 15.

### File name extension

Select from all file name extensions that exist within the info set.

At a maximum, the 20 top-ranked values are displayed as facet filters although there might be more.

### Object size

Select from up to six size ranges.

### Created date

Select from up to four date ranges.

## Modified date

Select from up to four date ranges.

## Reviewer

Select from the available set of reviewers.

This facet is available only if the result set contains any documents with a reviewer tag.

## Error code

Select from the processing errors that are returned for the info set. The facet values have the format *error\_message* (*type* *error*) where *type* indicates the processing step where the error occurred (Content, Binary, System).

Example: Unknown file format (Content error)

At a maximum, the 20 top-ranked values for each processing step are displayed as facet filters although there might be more. In contrast to the other facets, the sequence of the facet values is not necessarily determined by the frequency of occurrence. The values are grouped by type, but the order in which the types are listed is not predefined. Within a type, top-most values show up first.

## Processing state

Select from up to 12 processing states. The facet values have the format *type* - *state* where *type* indicates the processing step where the error occurred and *state* the state of the individual processing step (Not performed, Successful, Partial, Failed).

Example: Content - Successful

## Volume

Select from the source volumes that contribute to the current info set.

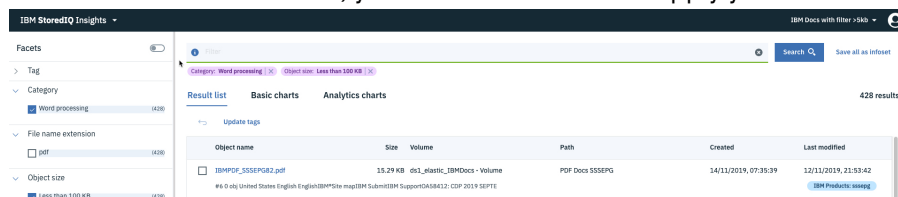
At a maximum, the 20 top-ranked values are displayed as facet filters although there might be more.

## Path

Select from the file paths found in the info set.

At a maximum, the 20 top-ranked values are displayed as facet filters although there might be more.

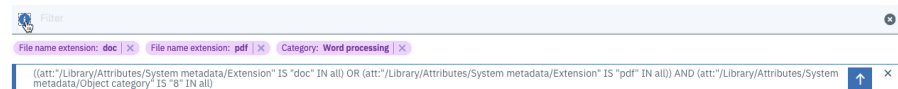
For each selected facet filter, a search tag is added to the breadcrumb trail beneath the filter box so that you can see your search criteria at a glance. By default, search results are instantly displayed when you select a facet filter. If you want to select multiple facet filters for a single search, you can use the toggle switch in the **Facets** pane to turn off automatic search. Alternatively, you can use the standard key-click combinations for multi-select for your operating system, which also turns off automatic search. With automatic search turned off, you must click **Search** to apply your filter criteria.



You can further refine the results by either adding or removing facet filters.

Instead of selecting facet filters, you can create your own queries. Also, you can combine a manual query with one or more facet filters. Queries that you build manually must follow the IBM StoredIQ filter syntax. For more information, see [“Filter condition reference and expression language”](#) on page 16.

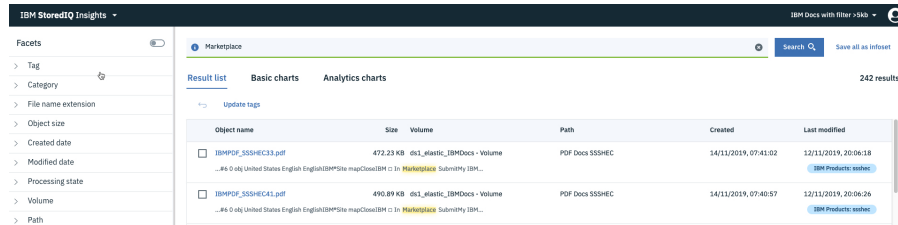
When you combine facet filters and other search terms, you might want to view and maybe even edit the composed query. In this case, click the information icon next to the filter box to see the complete query. You can transfer the composed query to the filter box for editing by clicking the up arrow.





## What do the results look like?

Objects that match the specified criteria are displayed instantly in the result list in the main content area. Highlighting of search hits and the option to preview document content and to view document metadata directly help you evaluate the results.



By clicking an object name, you can access the object's content preview (if available) and its metadata. You can switch between an HTML preview of the content and the object's metadata.

If the administrator created tags, you can also manage tags for one or more data objects from the result list: add, update, or remove tags. For more information about tags, see [“Document tagging” on page 6](#).

If the reviewer capability is in use, you can assign, change, or remove reviewers from data objects. For more information about the reviewer capability, see [“Assigning reviewers” on page 10](#).

You can save any result list as an infoset, which then becomes also available in IBM StoredIQ Data Workbench as a regular user infoset.

The basic and advanced charts provide visual representations of your search results.

## Searching and evaluating results

Search your data with IBM StoredIQ Insights for fast access to relevant information.

Remember that only data managed by data servers of the type DataServer - Distributed is searchable by IBM StoredIQ Insights. Therefore, you might see different object counts for mixed infosets in an environment with both types of data server (DataServer - Distributed and DataServer - Classic). The overall list shows the total objects count, which includes data managed by classic data servers. After you select an infoset, only the data managed by distributed data servers is reflected. Object counts can also differ when you refine a mixed infoset and save the result set as a new infoset. For **Insights Reviewer** users, the object count usually is different because their result lists contain only those data objects that are tagged with the respective reviewer tag.

### 1. In IBM StoredIQ Insights, select an infoset.

The title bar shows your current selection. You can switch infosets by selecting a different one from the list.

You might notice that the number of data objects that is shown for the selected infoset is different from the number that is shown in the overall list. Such mismatching object counts can occur in these cases:

- When the infoset contains data from several volumes some of which cannot be searched by IBM StoredIQ Insights (where the content is not managed by a distributed data server)
- For an **Insights Reviewer** user because only a subset of data objects is assigned for review (that is, tagged with the respective reviewer tag)

### 2. Enter your search criteria and narrow the search results to find particular documents of interest:

- Drill down by selecting facet filters one by one, where each facet filter is applied to the results of the preceding search.
- Apply several facet filters at once. Set the filter mode to manual and select facet filters as required. Alternatively, use the standard key-click combination for your operating system. Then, click **Search**.
- Enter a query in the filter field and click **Search**.

The query must follow the IBM StoredIQ filter syntax. For more information, see [“Filter condition reference and expression language”](#) on page 16.

- Combine facet filters and manual queries to create complex search criteria.

To remove a facet filter, either clear the respective checkbox in the **Facets** list or click the **X** on the respective label in the breadcrumb trail.

3. Check the search results.

Depending on your query, search terms are highlighted. To access a document's content preview and metadata information, click the object name in the result list. The object type might not allow for a content preview. In this case, you see only the metadata information.

4. Refine the results by repeating steps [“2” on page 5](#) and [“3” on page 6](#) until the results meet your needs.

5. Save the result set as an info set.

The info set then becomes also available in IBM StoredIQ Data Workbench as a regular user info set.

If your base info set is a mixed info set, the filter that you created is internally applied to all data objects in the info set. Therefore, the new info set can also be a mixed info set and might contain more objects than indicated in the IBM StoredIQ Insights user interface.

## Query hints and tips

Find some tips for avoiding The query expression parser returned an error. messages.

Usually, this error message results from special characters in your search string. Some of these characters serve a special purpose and you must enclose them in single or double quotes if you want to search for such a character literally. These special characters generate errors unless surrounded by quotation marks:

```
@ / & : ! ( ) { } [ ] ^ ~ , # =
```

A colon (:) without quotation marks is allowed in att:, cc:, ia:, re:, and tag: expressions.

Braces, brackets, and parentheses usually come in pairs (opening and closing) and are used, for example, for grouping items or with full-text macros.

## Document tagging

Tags are keywords that can be assigned to data objects in an info set. Data experts use them to label and categorize content so that they can easily cull relevant and related information.

In IBM StoredIQ Insights, tags are organized in high-level categories called *tag groups* for easier management. Except for the **Reviewer** tag group, an administrator must create and enable tag groups and tags before data experts can use them. Examples of such tag groups and the tags that they contain are these ones:

- Tag group **Sensitivity** to indicate the data sensitivity level. The set of tags might be **Ultra, Highly, Moderately, Normal, None**.
- Tag group **Classified** to indicate the security level of sensitive documents. The set of tags might be **Top Secret (TS), Secret, Confidential, Restricted**.

The **Reviewer** tag group and its tags are maintained by the system.

The administrator must also create any tag groups and tags that are used in automatic document classification with a CDA model. Enable such tag groups and tags only if you want to permit manual removal of automatically assigned tags.

Tag-related interface elements show up in the IBM StoredIQ Insights user interface only after at least one tag group was enabled. Data experts can then assign one tag per enabled tag group to a data object. Such tags become available as facet filters so that they can be used in queries in IBM StoredIQ Insights.

Reviewer tags are grouped in a separate facet.

## Administering tag groups and tags

Create and maintain tag groups and tags that data experts can use to label and categorize documents.

You must have IBM StoredIQ Insights administration rights to be able to complete this task, which means you must have the **Admin** and **Insights User** roles assigned.

Before you start to add tag groups and tags, think carefully about which tags you might need and how to group them best. Establish a tag system to ensure that data experts can categorize content effectively. The more thoroughly thought through your tag system is, the easier it is to find the content again later. While you can create as many tag groups with as many tags as you like, consider that a vast number can become unmanageable quickly. Also, consider that you cannot delete any tags or tag groups. You can only disable tags or tag groups that you don't want to be used.

No default tag groups are provided with the product. You must create all definitions from scratch. This also applies to tag groups that are used with CDA models.

To make tag groups and tags available to data experts, you must explicitly enable them. If a tag or tag group is no longer to be used, disable it. Existing tagging is not removed when you disable a tag. Thus, even disabled tags can show up in the facet pane.

Tag groups and tags are public. All data experts can use the same set of tags.

The exception to these rules is the **Reviewer** tag group, which is created automatically. The **Reviewer** tag group is maintained by the system. Whenever a user is assigned the **Insights Reviewer** role, a reviewer tag is created for this user. If this role is removed from the user's profile or if the user account is deleted, the respective reviewer tag is disabled and the person can no longer be assigned as reviewer. However, disabled reviewer tags are not removed and still show up in search results. Because the **Reviewer** tag group is system managed, it does not surface in the list of tag groups in **Insights administration**. Thus, you cannot edit, disable, or delete it.

Some restrictions for naming tag groups and tags apply:

- Names must not start with \$SIQ\$ in any combination of upper- and lowercase letters.
- Names must not contain a colon (:).
- Starting with this release, tag names and tag group names are limited to 30 characters.

1. In IBM StoredIQ Insights, click the user icon and select **Insights administration**.

2. To add a group, click **Add tag group** in the **Tag Groups** window, provide the following information, and save the settings. Otherwise, skip this step.

- A unique group name. The group name is added as a prefix to the tag names.
- A detailed description of the purpose of this tag group. Specifying a description is optional but consider that a detailed description helps users pick the proper tags for their documents.
- To make the tag group available for use, select the **Enabled** checkbox.

The new group is added to list of tag groups.

3. To update a tag group, edit it by clicking its name.

You can update the details at any time. When you disable a tag group after it was already used, existing tagging is persisted but the tag group is no longer available to the users.

a) To manage tag definitions, switch to the **Tags** tab.

- To add a tag, click **Add tag**. Then, provide a name, a description, and chose whether to enable the tag for use. The tag name must be unique within the tag group. A detailed description helps users pick the proper tag. If the tag group is not enabled, the setting of the **Enabled** option is irrelevant.
- To update a tag definition, click the tag name. When you update tag details, the same considerations apply as for updating tag group details.

b) Save any changes.

Any enabled tags within an enabled tag group are now ready for use.

# Assigning tags

Use tags to label and categorize data objects. You can use these tags for advanced searching and for managing your data.

One or more sets of enabled tag groups each with at least one enabled tag must be available.

You can add tags to documents and you can manage existing tags from the result list. You can combine tags from several tag groups to create fine-grained classifications for documents.

Tag assignments are shown in the result list. If a data object has more than one tag assigned, the result list entry shows **x tags applied**. Click this control to view the list of assigned tags.

Reviewer tags are a special kind of tags. For more information about working with these tags, see [“Assigning reviewers” on page 10](#).

Tags can also be assigned automatically by applying a CDA model cartridge to an info set. The behavior of such tags is slightly different. You can identify such a tag by hovering over it. The hover help does not only show the tag group and tag name but also the name of the cartridge that set the tag and the confidence level for the tag setting.

1. From the result list, select one or more documents.

You might want to check a document's content and metadata before you add, update, or remove tags. In this case, click its name to open the document information window where also an **Update tags** button is available.

2. Click **Update tags** to display the list of tag groups.

You can choose whether you want to update just the selected set of documents or all documents in the result list.

Before you select a tag group, you might want to check its description. To see the description, click the information icon next to tag group name.

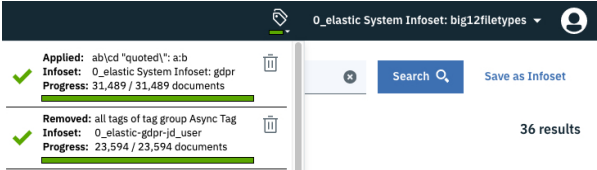
3. Click a tag group to display the available tags.

To see the description of a tag, click the information icon next to its name. A highlighted tag indicates that this tag is already assigned to the document. If you selected more than one document to work with, the overall number of documents and the number of documents that have this specific tag assigned is also shown.

4. Assign a tag.

You can assign one tag per tag group to a document.

Option	Description
<b>Apply to a selected set of documents</b>	<p>Click the respective tag:</p> <ul style="list-style-type: none"><li>• If it is assigned to all or part of the selected documents, the tag is removed from those documents to which was assigned. Documents that don't have this tag assigned remain unchanged.</li><li>• If it is not assigned to any of the selected documents, the tag is added to all documents. If any documents have tags from this tag group assigned, those are removed.</li></ul> <p><b>Tags set by a CDA model cartridge:</b> If the respective tag group is disabled, such tags cannot be removed.</p> <p>If the respective tag group is enabled, an enabled tag can be explicitly removed. A disabled tag cannot be explicitly removed but can be</p>

Option	Description
	overwritten by selecting another tag from this group.
<b>Apply to all documents in the result list</b>	<p>Click the respective tag:</p> <ul style="list-style-type: none"> <li>• If the tag is assigned to none or part of the documents, the tag is added to all documents. If any documents have tags from this tag group assigned, those are removed.</li> <li>• If the tag is applied to all documents, it is removed from all documents.</li> </ul> <p><b>Tags set by a CDA model cartridge:</b> If the respective tag group is disabled, such tags cannot be removed.</p> <p>If the respective tag group is enabled, an enabled tag can be explicitly removed. A disabled tag cannot be explicitly removed but can be overwritten by selecting another tag from this group.</p> <p>The update task might be moved to the background depending on the server response. When a tag update task is initiated, a task ID is returned that is used for querying the task status. If on the first progress query the task is still in progress, the task continues to run in the background. In this case, progress information is added to the title bar.</p>  <p>The progress bar shows the overall progress for all tag update tasks that you initiated. You can access information about the individual tag update tasks from there. The information that you see in the task list for each task includes which tag is applied, tags from which tag group are removed, the info set name, and progress information. In addition, you get information about any applied filter, and the start and completion times when you hover over a task entry.</p> <p>If any of the update tasks fails, the color of the progress bar changes to red. In the task list, failing tasks are indicated by a warning icon.</p> <p>The task list is maintained beyond logout. It is retrieved when you log back in so that you can check the status of any tasks that you started during your previous session. You can remove tasks manually from the list by clicking the trash</p>

Option	Description
	<p>icon. Tasks are removed automatically 7 days after completion or when the gateway server is restarted.</p> <p>You cannot pause or cancel update tasks that run in the background. The information in the UI is not automatically updated when an asynchronous update task is complete. You might need to refresh your browser to see the results.</p>

You can remove tags from individual documents by clicking the **X** on the respective tag in the result list entry. Tags that were set by a machine learning model do not have the **X** so that they cannot be removed by mistake.

Assigned tags become available as facet filters so that you can use them in your queries.

## Assigning reviewers

Distribute review tasks and restrict access to data objects by assigning a reviewer.

You must be logged in as an administrator or as a regular IBM StoredIQ Insights user.

At least one user account with the **Insights Reviewer** role must exist in IBM StoredIQ.

Provide an info set with data objects for review by a specific user. Only one user at a time can be set as reviewer.

Reviewer tags are maintained by the system. For each user with the **Insights Reviewer** role, a reviewer tag with the user's first, last, and user name exists in the **Reviewer** tag group. If the role is removed from the user's profile or if the user account is deleted, the respective reviewer tag is disabled. Reviewer assignments show up in the result list like any other tags but are labeled differently for easier distinction:

- General tags have labels in this format: **taggroup\_name:tag\_name**
- Reviewer tags have labels in this format: **first\_name last\_name (username)**

For searching, reviewer tags are grouped in a separate **Reviewer** facet.

1. Select an info set, filter as required for the review task, and create a new info set.

You can also assign reviewers without creating new info sets but it's easier to track the review tasks if you do.

2. Open the newly created info set and click **Update tags** to display the list of tag groups.

3. Select to update all documents in the result list.

4. Click the **Reviewer** tag group to display the available reviewers.

A highlighted entry indicates that this reviewer is already assigned to the document. If you selected more than one document to work with, the overall number of documents and the number of documents that have this reviewer assigned is also shown.

5. Assign a reviewer.

Only one reviewer can be set at a time. You have several options:

Option	Description
<b>Apply to a selected set of documents</b>	<p>Click the respective reviewer tag:</p> <ul style="list-style-type: none"> <li>• If the reviewer was assigned to all or part of the selected documents, the reviewer is removed from the respective documents. Documents that didn't have this reviewer assigned remain unchanged.</li> </ul>

Option	Description
	<ul style="list-style-type: none"> <li>If the reviewer was not assigned to any of the selected documents, the reviewer is set on all documents. If any documents had other reviewers assigned, those are removed.</li> </ul>
<b>Apply to all documents in the result list</b>	<p>Click the respective tag:</p> <ul style="list-style-type: none"> <li>If the reviewer was assigned to none or part of the documents, the reviewer is set on all documents. If any documents had other reviewers assigned, those are removed.</li> <li>If the reviewer was assigned to all documents, the assignment is removed from all documents.</li> </ul> <p>If the update task is running in the background, the task list basically has the same information as for regular tags. However, the <b>Applied</b> and <b>Removed</b> fields list the reviewer tag. For more information about asynchronous tagging and the task progress list, see the topic about <a href="#">assigning tags</a>.</p>

You can remove reviewers from individual documents by clicking the **X** on the respective reviewer tag in the result list entry.

The reviewer has access to the assigned data objects but no others and can work with the content, refine the info set, apply facets, create charts, and, if the role assignment allows for it, create new info sets. In addition, the reviewer can assign tags as required to approve content so that other users can work with the results. A reviewer can assign a different reviewer tag to delegate the review of a specific data object to another user or remove the reviewer tag altogether. In either case, the reviewer loses access to the respective data object.

## Visualizing your data

Visualizing your search results in graphical ways provides further insights into your data. Looking at data from different perspectives can help you understand large amounts of data quickly.

IBM StoredIQ Insights provides different chart types for different ways to communicate data. Each visualization initially represents your search results showing the distribution of documents across facets and can be refined.

### Basic charts

Use the basic charts to compare facets with a one-on-one relation. Indexed annotations or tags cannot be used as dimensions in basic charts because documents can be associated with more than one annotation or tag:

- Treemap charts give a hierarchical view of the distribution of the data objects in your result set across the facets that you select as the primary and optionally as secondary dimension. Treemap charts use differently colored nested tiles to represent data. The size of each tile is proportional to the number of hits for the facet filter it represents. Thus, you can get an overview of contributions.
- In bar charts, you can compare the contributions of the facet filters. Each bar represents the objects that were found for a specific facet filter.

In a stacked bar chart, the bars are subdivided into multiple sections, each with a different color. Each section shows its relative contribution to the entire bar.

In a clustered bar chart, the bars are arranged in groups so that you can compare values side by side. Each bar has a different color.

## Analytics charts

Use the analytics charts to see the distribution of indexed annotations or tags across a specific facet. Analytics charts combine stacked bar charts and clustered bar charts to visualize the information.

## Creating basic charts

Use basic charts to visualize the distribution of document across facets other than indexed annotations and tags.

Available chart types are treemap chart and bar chart. If you select both a primary and a secondary dimension, you can also choose between a stacked bar chart and a clustered bar chart.

A chart initially represents the current result set. It is automatically updated when you change the filter criteria.

Each facet has a certain limit of facet filters that can be displayed in the user interface although the info set might contain more values. For more information about the limit for each facet, see the [facet descriptions](#). Findings that are not listed in the facet pane or named in the legend are still reflected in the chart.

- Treemap chart: such values are indicated by a tile in gray color, dark gray for the primary dimension and light gray for the secondary dimension. The tile is labeled **other dimension**.
- Bar chart (one dimension): such values are indicated by an entry at the bottom of the y-axis named **other dimension**.
- Stacked or clustered bar chart (two dimensions): for the primary dimension, such values are indicated by an entry at the bottom of the y-axis named **other dimension**. The legend for this bar shows **unknown**, which means the secondary dimension can have any value. For all other primary dimension values, secondary dimension values that are outside the range of values that can be listed show in the legend as **other**.

For treemap charts, a navigation path is created as you click entries in the charts.

1. To create a basic chart from the information in the result list, click the **Basic charts** tab.

The default settings for a basic chart are as follows:

- The primary dimension is set to **Volume**.
- No secondary dimension is set (**(none)**).
- The chart type is set to **Treemap chart**.

2. To create different charts, you have these options:

Chart type	Options
<b>Treemap chart</b>	<ul style="list-style-type: none"> <li>• Change the primary dimension. In this case, the secondary dimension is reset to <b>(none)</b> if you selected one before.</li> <li>• Add a secondary dimension.</li> <li>• Update your filter criteria by applying further facet filters or entering a query.</li> </ul> <p>You can combine these options as required.</p>
<b>Bar chart</b>	<ul style="list-style-type: none"> <li>• Change the primary dimension. In this case, the secondary dimension is reset to <b>(none)</b> if you selected one before.</li> <li>• Update your filter criteria by applying further facet filters or entering a query.</li> </ul>



Chart type	Options
	<p>You can combine these options as required.</p> <p><b>Hint:</b> When you add a secondary dimension, the chart type changes to <b>Stacked bar chart</b>.</p>
<b>Stacked bar chart</b>	<p>Stacked bar charts are available only if a secondary dimension is selected. In this case, the chart includes a legend that corresponds to the facet filters that are displayed in the facet pane.</p> <ul style="list-style-type: none"> <li>• Change the primary dimension. In this case, the secondary dimension is reset to <b>(none)</b> if you selected one before.</li> <li>• Change the secondary dimension.</li> <li>• Update your filter criteria by applying further facet filters or entering a query.</li> <li>• For the secondary dimension, you can hide or show the results for any facet filter by clicking its entry in the legend. This does not change the result set.</li> </ul> <p>You can combine these options as required.</p> <p><b>Hint:</b> When you remove the secondary dimension, the chart type changes to <b>Bar chart</b>.</p>
<b>Clustered bar chart</b>	<p>Clustered bar charts are available only if a secondary dimension is selected.</p> <ul style="list-style-type: none"> <li>• Change the primary dimension. In this case, the secondary dimension is reset to <b>(none)</b> if you selected one before.</li> <li>• Change the secondary dimension.</li> <li>• Update your filter criteria by applying further facet filters or entering a query.</li> <li>• For the secondary dimension, you can hide or show the results for any facet filter by clicking its entry in the legend. This does not change the result set.</li> </ul> <p>You can combine these options as required.</p> <p><b>Hint:</b> When you remove the secondary dimension, the chart type changes to <b>Bar chart</b>.</p>

Depending on the amount of data, bar charts might not fit the window size. You can either scroll down to view all data or resize the chart. To do so, drag one of the resizing handles on the scrollbar thumb until the chart is the right size.

3. Optional: To see additional information, hover over a tile or bar in your chart.

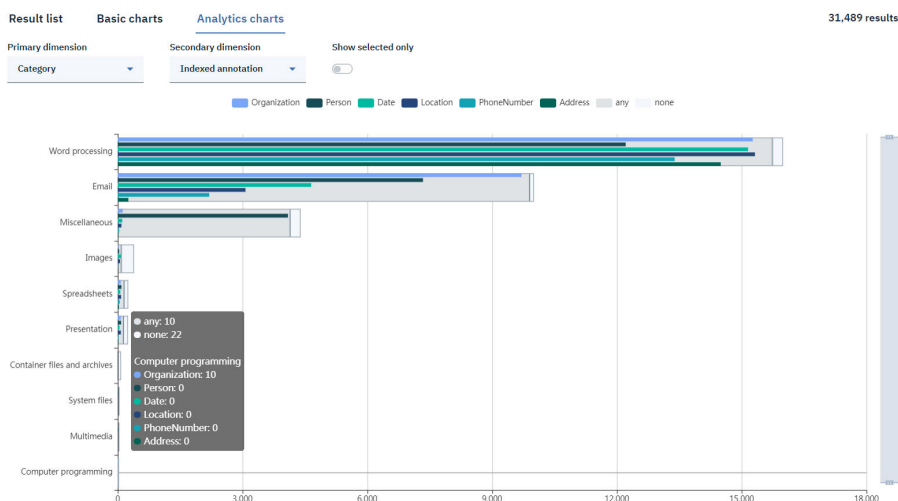
## Creating analytics charts

Use analytics charts to see how many documents of the result set contain indexed annotations or tags per the selected primary dimension facet.

An analytics graph illustrates the following information:

- How many documents per selected primary dimension facet contain at least one indexed annotation or tag (dark gray stacked column).
- How many documents per selected primary dimension facet do not contain any indexed annotation or tag (light gray stacked column).
- The true relative proportions of the amount of data per primary dimension that contains at least one indexed annotation or tag (relative height of the stacked column in dark and light gray).
- The absolute and relative contributions of the individual facet filters (clustered columns in different colors).

A chart initially includes all indexed annotations or tags that were found in the result set, no matter what you selected from these facets in the facet filter. Reviewer tags are just a different type of tags and are therefore reflected within the **Tags** dimension.



Each facet has a certain limit of facet filters that can be displayed in the user interface although the infocset might contain more values. For more information about the limit for each facet, see [the facet descriptions](#). Findings that are not listed in the facet pane or named in the legend are still reflected in the chart.

For the primary dimension, such values are indicated by an entry at the bottom of the y-axis named **other dimension**. The legend for this bar shows **unknown**, which means the secondary dimension can have any value. For all other primary dimension values, indexed annotations or tags that are outside the range of values that can be displayed, have an entry that is listed in the legend as **other**.

1. To create an analytics chart from the information in the result list, click the **Analytics charts** tab.

By default, the chart shows all indexed annotations that were found in the result set with the primary dimension set to **Volume**.

2. To create other charts, you have these options:

Secondary dimension	Options
Indexed annotation	<ul style="list-style-type: none"> <li>• Change the primary dimension.</li> <li>• Update your filter criteria by applying further facet filters or entering a query.</li> <li>• Set <b>Show selected only</b> to on to show only results that contain exactly those indexed annotations that you selected in the facet pane.</li> <li>• Hide or show the results for any indexed annotation by clicking its entry in the legend. This does not change the result set.</li> </ul> <p>You can combine these options as required.</p>

Secondary dimension	Options
Tag	<ul style="list-style-type: none"> <li>• Change the primary dimension.</li> <li>• Update your filter criteria by applying further facet filters or entering a query.</li> <li>• Set <b>Show selected only</b> to on to show only results that contain exactly those tags that you selected in the facet pane.</li> <li>• Hide or show the results for any tag by clicking its entry in the legend. This does not change the result set.</li> </ul> <p>You can combine these options as required.</p>

Depending on the amount of data, the chart might not fit the window size. You can either scroll down to view all data or resize the chart. To do so, drag one of the resizing handles on the scrollbar thumb until the chart is the right size.

- Optional: To see additional information, hover over a bar in your chart.

---

# Filter condition reference and expression language

The following topics provide in-depth information about full-text and attribute conditions and how to use the expression language.

## Full-text syntax

---

The full-text editor supports full-text syntax.

The full-text feature supports a wide range of full-text syntax to broaden or narrow filters. A one-word term like `September` returns results from all harvested data objects that include the term in their content. Word matching is case-insensitive. A list of words matches any of those words. For example, `September OR October` match all objects that contain either September or October.

### Term expansion

When lemmas are turned off (by default), a simple term such as `went` returns only those documents that contain the exact same character sequence. For more information, see the section on configuring full-text index settings in the IBM StoredIQ Data Server Administration documentation.

When lemmas are turned on, the simple term `went` is automatically expanded to return documents that contain words with the same meaning as `went`, such as `go`, `goes`, `going`, and `gone`.

Automatic term expansion applies to any terms. Term expansion can be disabled by placing the term within single quotation marks. Term expansion does not occur when a wildcard is present.

**Note:** To use term expansion for the anticipated results, lemmas must be turned on before the documents are indexed.

### Grouping

When you are examining text, IBM StoredIQ Data Workbench evaluates the requested operators and gives precedence to some over others. These rules are described in the following list.

When parentheses are used to group operators, the expression in parentheses is given the highest precedence. If no parentheses are used in the expression, associativity rules are used instead. Right-associative means that the operator is grouped initially with the operand to its right. Left-associative means that the operators and operands to the left are evaluated before the operators and operands to the right.

- The NOT operator is right-associative.

`"October NOT November OR December"` means `"(October AND (NOT November)) OR December"`.

- Proximity operators (`w/` and `pre/`) are left-associative.

`"October w/3 November pre/5 December"` means `"(October w/3 November) pre/5 December"`, where `"October"` and `"November"` must be within three words of each other, but both must be before `"December"` by no more than five words.

`"October w/3 November and December"` means `"(October w/3 November) and December"`.

- The AND operator is left-associative.

`"October AND November OR December"` means `"(October AND November) OR December"`.

- The OR operator is left-associative. If you do not specify an operator, expressions are ORed.

`"October November OR December"` is equivalent to `"(October OR November) OR December"`.

## Boolean Operators

This table lists Boolean operator syntax and descriptions.

Table 1. Boolean operator syntax	
Item	Description
AND &&	<ul style="list-style-type: none"><li>When used between two terms, the AND operator ensures that both terms appear in the data object. For example, the filter <code>chairman AND stock</code> returns only data objects where <code>chairman</code> and <code>stock</code> both appear.</li><li>The double ampersand (&amp;&amp;) can be used instead of the AND operator. AND can be uppercase or lowercase.</li></ul>
OR 	<ul style="list-style-type: none"><li>When used between two terms, the OR operator returns data objects where either term appears. For example, the filter <code>chairman OR stock</code> returns data objects that have <code>chairman</code>, <code>stock</code>, or both in the data object.</li><li>The double pipe (  ) can be used instead of the OR operator. OR can be uppercase or lowercase.</li></ul>
NOT !	<ul style="list-style-type: none"><li>When used before a term, the NOT operator ensures that the following term does not appear in any data objects returned. For example, <code>Apple NOT IBM</code> returns data objects where word <code>Apple</code> appears, but not <code>IBM</code>. You can think of the NOT item as "and not" or "but not".</li><li>The exclamation point (!) can be used instead of the NOT operator. NOT can be uppercase or lowercase.</li></ul>

## Multiple Terms or Term Phrase

This table lists multiple terms or term phrase syntax items and descriptions.

Table 2. Terms	
Item	Description
" "	<ul style="list-style-type: none"><li>Double quotation marks around a term phrase return data objects that have that phrase in order in a data object. For example, <code>"September 20 1998"</code> returns only data objects with the exact phrase of <code>September 20 1998</code>. Double quotation marks allow for wildcards in terms whereas single quotation marks do not.</li><li>Stop words are removed from a double-quoted term phrase. For example, <code>stock and option</code> returns documents with <code>stock option</code>.</li><li>Asterisks (*) and question marks (?) are treated as wildcards. For example, <code>stop option*</code> returns both <code>stock option</code> and <code>stock options</code>.</li><li>Automatic term expansion is performed if lemmas are turned on. For example, <code>fox jumps</code> returns any of these phrases: <code>fox jumps</code>, <code>fox jumped</code>, <code>foxes jump</code>, and so on.</li></ul>
' '	<ul style="list-style-type: none"><li>Single quotation marks around a term phrase return data objects that have the exact same character sequence in a data object. For example, <code>'february 4 2013'</code> returns exactly <code>february 4 2013</code>.</li><li>Asterisks (*) and question marks (?) are treated as regular characters. For example, <code>'*ism'</code> returns exactly <code>*ism</code>.</li><li>Stop words are not removed. For example, <code>'to be or not to be'</code> returns exactly <code>to be or not to be</code>.</li><li>Term expansion is not performed even if lemmas are turned on. For example, <code>'went'</code> returns exactly <code>went</code>.</li></ul>

Table 2. Terms (continued)	
Item	Description
Multi-term, no Boolean operators	A filter for a multiple word without Boolean operators or quotation marks is an implied OR filter. For example, September 20 1998 returns data objects that have September OR 20 OR 1998.
Grouping	<ul style="list-style-type: none"> <li>Grouping with Boolean operators and parentheses allows users to narrow results by using attributes. For example, (\$10 OR billion) AND options returns data objects where \$10 or billion might be in the data object, but options must exist.</li> <li>Boolean operators (AND, OR, NOT) work in a grouping or in a proximity group. So ("Sahara Desert" AND "Southeast Asia") OR ("Pacific Ocean" AND "Bavaria") returns data objects that have Sahara Desert and Southeast Asia or Pacific Ocean and Bavaria.</li> </ul>

## Wildcards

This table lists wildcard syntax items and descriptions.

Table 3. Wildcard	
Item	Description
*	<ul style="list-style-type: none"> <li>A single asterisk finds multiple characters or letters at beginning of a term. An asterisk that is placed before a partial word finds terms that end with those letters. For example, a filter for *olol returns results for betaxolol, carteolol, levobetaxolol, levobunolol, metipranolol, and timolol.</li> <li>An asterisk matches any number of characters in the middle of a term. An asterisk that is placed in the middle of a term finds terms that have those characters on either side of the asterisk. For example, te*t returns results for text, test, tent, and tempest.</li> <li>An asterisk that is placed after a partial word finds terms that begin with those letters at the end of a term. For example, a filter for ap* returns results for Apple, April, applied, and application.</li> <li>An asterisk alone finds all objects.</li> </ul> <p><b>Note:</b> An asterisk within single quotation marks is treated as a regular character.</p>
?	<p>The question mark matches a single character. For example, compl? returns results for comply but not complies or compliance. A filter for ap? finds ape, but not apple. To filter for more than one character, use a question mark for each letter, for example, ap??e.</p> <p><b>Note:</b> A question mark within single quotation marks is treated as a regular character.</p>

## Proximity

This table lists proximity syntax items and descriptions.

Table 4. Proximity syntax

Item	Description
w/x	<ul style="list-style-type: none"> <li>Filtering uses the w/x syntax (x represents the number of terms between the two terms) to use the relationship between terms to get more targeted results. A two-term filter with one term that appears at the beginning of a data object and the other at the end might not be as relevant as when the terms appear in the same paragraph or sentence. For example, in an SEC data object, chairman w/15 options is a more relevant filter than chairman AND options appearing in the same data object that might be several pages long.</li> <li>w/0 matches terms that are next to each other in any order.</li> <li>w/1 matches for terms with one word between them. The terms can appear in any order in the document.</li> <li>Grouping with proximity allows for even more targeted results. For example, ("Sahara Desert" "Southeast Asia") w/5 ("Pacific Ocean" Bavaria) finds either Sahara Desert or Southeast Asia within five terms of Pacific Ocean or Bavaria.</li> <li>Terms can have wildcards or phrases.</li> </ul>
pre/x	<ul style="list-style-type: none"> <li>With the pre/x syntax (x represents the number of terms between the two terms), you can further target proximity filters by dictating which term comes before another term in the data object and how far before it is. For example, chairman pre/15 options returns data objects where chairman comes up to 15 terms before options.</li> <li>pre/0 matches terms that are next to each other and in the same order.</li> <li>pre/1 matches terms with at most one word between them and in the same order.</li> <li>Terms can have wildcards or phrases.</li> </ul>
w/p	The w/p syntax lets you find terms within the same paragraph. The terms can appear within the paragraph in any order.
pre/p	The pre/p syntax lets you find terms within the same paragraph. The terms must appear in the specified order.
w/s	The w/s syntax lets you find terms within the same sentence. The terms can appear within the sentence in any order.
pre/s	The pre/s syntax lets you find terms within the same sentence. The terms must appear in the specified order.

## Full-text macros

The IBM StoredIQ full-text filter supports a set of predefined macros.

A macro is an abbreviated set of instructions that are used in common filter patterns. For example, the filter {SSN} finds any document that contains Social Security numbers.

Full-text macros have the following characteristics:

- A macro is an alphanumeric name that is enclosed in curly braces, for example, {SSN}.
- A macro's name is case insensitive. For example, the macro {ssn} is the same as the macro {SSN}.
- The macro's value is replaced in the macro's full-text query. Then, the filter is parsed and run.
- Any unrecognized macro results in a syntax error.
- Regular expressions, that is named character classes, do not match macros. That is, a regular expression like \p{Lower} (matches any lowercase ASCII character) is not recognized as a macro.

The following macros are defined within the full-text feature of IBM StoredIQ. Each of these macros is a high-confidence macro. It means that the filter `{SSN}` finds any document that contains Social Security numbers with high confidence that the matches really are Social Security numbers. To find anything that might be a Social Security number, use the low-confidence macro `{SSN-1c}`.

Macros can be combined with any other full-text expression, such as `{SSN} w/5 "John Doe"`.

- `{SSN}`: Social Security numbers, which can include dashes and numbers
- `{USPhone}`: US phone numbers, which can include dashes, dots, parentheses, numbers, and letters. This macro can filter for mnemonics like 1-800-flowers.
- `{InternetAddress}`: Legal IP addresses, either as numbers or as names, which can include dots, numbers, and letters.
- `{EmailAddress}`: Legal email addresses, which can include @, dots, numbers, and letters.
- `{CCAmex}`: American Express credit-card numbers, which can include dashes and numbers.
- `{CCDiners}`: Diners Club credit-card numbers (including enRoute credit-card numbers), which can include dashes and numbers.
- `{CCDiscover}`: Discover credit-card numbers, which can include dashes and numbers.
- `{CCJCB}`: JCB credit-card numbers, which can include dashes and numbers.
- `{CCMC}`: MasterCard credit-card numbers, which can include dashes and numbers.
- `{CCVisa}`: Visa credit-card numbers, which can include dashes and numbers.
- `{CreditCard}`: Any credit-card number from the listed providers.

## Indexed annotations

---

To search for indexed annotations (terms indexed when a cartridge is applied), you can create filters with the query format `ia:filter_term`.

You can search by indexed annotations after the system administrator deployed the cartridge and created a Step-up Analytics action with the cartridge, and you ran the respective Step-up Analytics action on an infoSet. Set up an indexed annotation query by creating a full-text filter with a `ia:filter_term` expression.

Contact your system administrator for a list of deployed cartridges, their supported search results, and their associated actions. You can also find out about the supported results by checking the **Action Details** section of a selected Step-up Analytics action.

By default, the latest versions of the IBM provided cartridges that you can download from Fix Central support these indexed annotation terms:

### Basic cartridge

```
ia:BankAccountNumber
ia:EmailAddress
ia:IntraAddress
ia:IPAddress
ia:NationalIdNumber
ia:PhoneNumber
ia:TaxNumber
```

### Advanced cartridge

```
ia:Address
ia:Date
ia:GenderIndicator
ia:Location
ia:Organization
```



```
ia:Person
ia:PhoneNumber
```

The searches are case-sensitive. If the supported language field is populated, only documents in those languages are searched.

These cartridges can be modified to produce different sets of indexed annotations. For more information, see the documentation that is packaged with the cartridges.

## Regular expressions

---

To search for patterns, you can create full-text filters with regular expressions. The format for a regular expression query is `re: "regex"`.

For DataServer - Distributed, `re:` queries basically follow the Lucene regular expression syntax as described here:

[Elasticsearch Reference: Regular expression syntax](#)  
[Lucene RegExp class](#)

The character classes `\d`, `\D`, `\w`, and `\W` are supported, POSIX character classes such as `\p{Alpha}` or `\p{Alnum}` aren't. Complex constructs such as backreferences or lookahead are also not supported.

For DataServer - Classic, `re:` queries basically follow the Java regular expression syntax as described here: [Java regex pattern](#)

Patterns in IBM StoredIQ are always anchored at the beginning. For example, the string `abcde` would be a match for the query `re: "ab.*"` but not for the query `re: "bcd"`.

The IBM StoredIQ full-text index is a tokenized index. It does not include spaces or other special characters such as `/` `*` `\` `@` `\` `#` `"` and others. Therefore, you cannot create regular expressions that try to match spaces as part of a string.

For example, if you wanted to identify eight numbers in groups of four separated by a space like `1234 5678`, you would probably try the following regular expressions:

- `re: "\d{4} \d{4}"`
- `re: "\d{4}\s\d{4}"`
- `re: "\d{4}( )\d{4}"`
- `re: "\d{4}[ ]\d{4}"`

However, neither of these regular expression would yield a result. Instead, you can consider the string to be two 4-digit words. You could then use proximity search to join them:

```
re: "\d{4}" w/0 re: "\d{4}"
re: "\d{4}" pre/0 re: "\d{4}"
```

This type of query is called a *low confidence query* because it is more likely to produce false positives. In addition, such queries consume considerable processing time.

In contrast, indexed annotations from a regular expression cartridge can contain spaces and special characters because a cartridge is applied to the source text and not to the tokenized text in the index. So, depending on your use case, you might want to create a regular expression cartridge, run it on your data, and then use `ia:` queries instead of `re:` queries to find such strings.

## Extended ASCII characters

Extended ASCII characters are handled several different ways in the full-text filter. Some are special characters while others are treated as spaces. Some are treated as letters that depend on the spaces around the characters.

Special characters or punctuation are not indexed and cannot be searched for in filters. The following table describes what behavior applies to these characters in the filter.

Table 5. Extended ASCII Character syntax	
Extended ASCII Character Syntax	Description
Special characters	<p>Some special characters generate syntax errors because they cannot be processed. Either remove the character from the filter or surround the filter with single quotation marks. For example, a filter for <code>hello!</code> without quotations generates an error. However, if the filter is <code>'hello!'</code> (with single quotation marks), the query can be processed.</p> <p>In general, special characters are treated as a space in a filter and they are not indexed. For example, <code>Hello!</code> is indexed as <code>Hello</code>. Searching for <code>'Hello!'</code> returns zero object. To search for <code>'hello'</code>, data object that has the word <code>hello!</code> or <code>hello</code> is returned. However, there are some exceptions, and they are discussed later in this table</p> <p>Most special characters can also be searched for with quotation marks. For example, <code>"12/24/2014"</code> finds <code>12/24/2014</code>. The difference between quotation marks and single quotation marks is that quotation marks also have another function: to find words that are next to each other. Because of this, special characters that are used elsewhere in the syntax, such as <code>*</code> and <code>?</code>, are treated as the full-text operators for wildcards instead of as the characters <code>*</code> and <code>?</code>.</p> <p>These special characters generate errors unless surrounded by quotation marks: <code>@ / &amp; : ! ( ) { } [ ] ^ ~ , # =</code></p>
Period (.)	<p>A period with no space on either side is treated like a letter. For example, a filter for <code>172.16.0.46</code> finds the whole number of <code>172.16.0.46</code>, not partial numbers of <code>172</code>, <code>16</code>, <code>0</code>, <code>46</code>.</p> <p>A period with a space on either side is treated as a space and ignored. For example, a filter for <code>"1234 Mahana St."</code> ignores the period and just finds <code>1234 Mahana St</code> only without a period.</p>
At (@)	<p>An <code>@</code> symbol is indexed like a letter when there are no spaces on either side. For example, a filter for <code>jdoe@yourcompany.com</code> finds the entire email address. An <code>@</code> symbol with a space on either side or one side is treated as a space and ignored. For example, <code>@home</code> finds <code>home</code> in the data object.</p>
Hyphen (-)	<p>A hyphen is indexed like a letter when there are no spaces on either side. For example, a filter for <code>012-44-5668</code> finds the entire number, not <code>012</code>, <code>44</code>, <code>5668</code> as separate numbers. A hyphen with a space on either side is treated as a space and ignored.</p>
Slash (/)	<p>A slash is indexed like a letter when there are no spaces on either side. For example, a filter for <code>10/30/1940</code> finds the entire expression <code>10/30/1940</code> (and not <code>10 30 1940</code>) as separate numbers.</p>

Table 5. Extended ASCII Character syntax (continued)	
Extended ASCII Character Syntax	Description
Ampersand (&)	An ampersand is indexed like a letter when there are no spaces on either side. For example, a filter for S&P finds S&P in the data objects. However, a filter for Proctor & Gamble without quotation marks finds either Proctor or Gamble. A filter with quotation marks "Proctor & Gamble" finds the phrase Proctor Gamble only.
Colon (:)	A colon is considered a special character. It is treated as a space in filter and is not indexed. For example, 21:35:15 is indexed as three consecutive words, that is, 21, 35, and 15. Searching for '21:35:15' returns zero object. To search for '21 35 15', data object that has the word 21:35:15 or that has three consecutive words as 21, 35, 15 is returned. Specifying terms with colons without the single quotation marks results in an invalid syntax error.
Apostrophe (')	The apostrophe in the term is indexed as a whole word. For example, a filter for patient's returns results for patient's but not for patient. A filter for the name O'Hara returns results for O'Hara. An apostrophe is treated as a single quotation mark, so it must always be quoted with double quotation marks.
Comma (,)	Commas are left in when surrounded by numerals, and those numerals make a legitimate number. For example, "1,000" is indexed as 1,000. To search this term, specify 1,000 in single quotation marks or quotation marks. Otherwise, it results in an invalid syntax error. For the term "10,00", it is indexed as 10 and 00. To search for this term, specify "10 00".
Number sign (#), dollar sign (\$), percent (%), equal sign (=), underscore (_)	These characters are treated as white space in the filter. For example, a search for \$100 returns results for the number 100 whether 100 refers to dollars, a percentage, or a whole number. The term 'JAVA_HOME' is indexed as 'JAVA HOME'. To search for this term, specify JAVA HOME in quotation marks.

## Attribute condition reference

The following section explains how to type in a reference to an attribute condition.

If you always select an attribute from the custom filters or one of the attribute groups, IBM StoredIQ takes care of the attribute reference for you. If you do not select an attribute from the custom filters or one of the attribute groups, then you need to review and use this section.

A typed attribute reference is composed of several parts: `attribute-type path/attribute-name`. The `attribute-type` part gives the specific type of attribute to which path and `attribute-name` refer. This prefix is a short prefix of the attribute type. Attributes can have duplicate names, even within one folder.

Table 6. Attribute type prefix	
Attribute-type Prefix	Description
att: or @	This prefix is a system-defined attribute or user-defined entity.

path describes the folder path to the attribute. Generally, it is optional, but it is required if the attribute name alone is not unique. You do not need to type the full folder path to an attribute.

If an attribute name or folder path name has spaces or punctuation in it, then you must enclose the complete name in double or single quotation marks, not just a part. However, the forward slash (/)

between folders or before the attribute name does not require the use of quotation marks. Do not enclose the attribute type prefix in quotation marks.

## Specification of attribute conditions

---

Attributes, entity, tag, or query can be divided up in several ways.

- What type of values does the attribute have? Attributes can have values of number, string, enumeration, boolean, or date. An enumeration attribute is an attribute that has a limited set of possible values, which are shown in a list.

When an attribute is a compound attribute, each of its subattributes also has a value type: string, enumeration, boolean, or date. The subattributes can also be either sparse or dense. Currently, all subattributes are single-valued. Compound attributes can also be sparse or dense. (Currently all compound attributes are sparse.)

- Do all objects have a value for this attribute? If the objects all have a value, it is known as a dense attribute. If the objects do not have a value, it is known as a sparse attribute. All objects have an `Object name` value, so `Object name` is a dense attribute. Only some objects have a `CC` attribute value, so `CC` is a sparse attribute.
- Can an attribute have multiple values within one object? Each object has one name, so `Object name` is a single-valued attribute. There might be multiple `To` values on an email, so `To` is a multivalued attribute. Compound attributes can be single-valued or multivalued. A compound attribute never exists where all of its subattribute values are missing.

These condition operators are applicable to each attribute category.

**Note:** For illustration purposes, some operators are shown in uppercase, and some are shown in lowercase. The parser recognizes operators in either uppercase or lowercase.

### Conditions for string attributes

Tags are considered single-valued, sparse string attributes. They are not considered enumeration attributes, even though they have a limited set of possible values.

```
attribute = value or attribute is value
```

```
attribute != value or attribute is not value
```

```
attribute contains value
```

```
attribute does not contain value
```

```
attribute starts with value
```

```
attribute does not start with value
```

```
attribute ends with value
```

```
attribute does not end with value
```

```
attribute matches value
```

```
attribute does not match value
```

The value can be a single word, a double-quoted string, or a single-quoted string. For matches, the value is a regular expression within quotation marks.

## Conditions for enumeration attributes

Some attributes have a limited set of predefined values that they can match, which are called enumeration attributes.

In this case, only '=' and '!=' operators are supported.

```
attribute = value or attribute is value
```

```
attribute != value or attribute is not value
```

A list of enumeration attributes is as follows:

- Container type
- Is-part-of
- Box reason
- Encryption type
- System processing state
- Content processing state
- Binary processing state
- System error code
- Content error code
- Binary error code
- Message type
- Latest review status
- Object category

## Conditions for date attributes

The conditions for date attributes are listed as follows.

```
attribute = value or attribute is value
```

```
attribute != value or attribute is not value
```

```
attribute < value
```

```
attribute <= value
```

```
attribute > value
```

```
attribute >= value
```

```
attribute between value1 and value2
```

```
attribute not between value1 and value2
```

The value can have one of two forms: an explicit date or a relative date that is offset from today.

An explicit date is in the form of YYYY-MM-DD, for example, 2009-03-08.

- Other date formats are supported.
- Time is not supported.
- The dates are all interpreted as midnight of the listed date.

A relative date offset is specified as a positive number, followed by a required date units, and the ago or ahead keyword. The ago keyword is the default and means a date in the past. The ahead keyword means

a date in the future. For example, 3 months ago, 3 days ahead, or 3 weeks ahead are legal relative dates. The allowed date units are as follows:

- days
- weeks
- months
- years

This function is also supported:

- -3 days = 3 days ago
- +3 weeks = 3 weeks ahead

**Note:** Months are considered as 30 days; years are considered as 365 days.

- Normally, relative date offsets are in the past, but some dates can be in the future, such as a Retention expiration date. For these references, use the ahead keyword instead of ago.
- Relative date offsets can also be given as a positive or negative number, such as -3 days or +3 months without the ago or ahead keyword. In this case, negative numbers are considered in the past and positive numbers are considered in the future.

## Conditions for number attributes

Conditions for number attributes are listed as follows.

```
attribute = value or attribute is value
```

```
attribute != value or attribute is not value
```

```
attribute < value
```

```
attribute <= value
```

```
attribute > value
```

```
attribute >= value
```

```
attribute between value1 and value2
```

```
attribute not between value1 and value2
```

The value is a positive or negative integer. Size attributes (Size and Size (on disk)) can take an optional size unit after the number. These amounts are the legal size units.

- b or bytes for single bytes
- k or kb for kilobytes
- m or mb for megabytes
- g or gb for gigabytes
- k or kb for kilobytes

**Note:** The default size unit is bytes. Size units are based on multiples of 1024, not 1000.

Event duration can have time units. These time amounts are the legal size units:

- m or minutes for minutes
- h or hours for hours
- d or days for days

## Conditions for sparse attributes

The conditions for multivalued and single valued sparse attributes are listed as follows.

```
attribute APPEARS  
attribute NOT APPEARS or attribute DOES NOT APPEAR
```

These conditions specify whether the attribute has any values for matching objects. In this case, it means that one of the subattributes has a value on this object. Tags are considered single-valued, sparse string attributes.

It is not an error to specify these conditions on dense attributes. However, if you do, the parser warns you that the condition is always `true` or `false`. These operators can also apply to compound attributes. In this case, it means that one of the subattributes has a value on this object.

**Note:** As a backward compatible measure, if `att: attribute appears` is given on a dense, boolean attribute, it is converted into `att: attribute = true`. If `not appears` is given, it is converted into `att: attribute = false`.

## Conditions for compound attributes

Conditions for compound attributes are listed as follows.

```
attribute WHERE <sub-attribute-condition>  
attribute WHERE (<sub-attribute expression>)
```

A `<sub-attribute-condition>` is just like any of the attribute conditions that are described except it is on a subattribute of the compound attribute. Additionally, a `<sub-attribute-condition>` can be just the subattribute name. It is equivalent to the subattribute **APPEARS**.

A `<sub-attribute-expression>` is any combination of `<sub-attribute-condition>`'s combined with **AND**, **OR**, or **NOT**. This expression must be enclosed in parentheses. The `<sub-attribute-expression>` can also be empty. In this case, it is the same as attribute **APPEARS**.

This example uses **WHERE** expressions:

```
att: "Site owner (SharePoint)" WHERE name = "John Smith"  
att: "Site owner (SharePoint)" WHERE (name = "John Smith" AND url APPEARS )  
att: "Site owner (SharePoint)" WHERE (name = "John Smith" OR url NOT APPEARS)  
att: "Site owner (SharePoint)" WHERE ()  
att: "Collaborator (Box)" WHERE "Collaborator Role" = "editor"
```

**Note:** This expression is the same as `att: "Site owner (SharePoint)" APPEARS`.

```
att: "Site owner (SharePoint)" WHERE (name)
```

**Note:** This expression is the same as `att: "Site owner (SharePoint)" WHERE name APPEARS`.

## Conditions for multivalued attributes

Conditions for multivalued sparse or dense attributes are listed as follows.

```
attribute APPEARS operator number
```

```
attribute APPEARS UNIQUE operator number
```

These operators can also apply to compound attributes, and the `operator` can be one of the following operators:

- `= >`
- `>=`
- `<`
- `<=`

The `operator number` specifies a condition on the number of values present for the attribute on any matching object.

- If `operator number` is not given, then `> 0` is assumed.
- If `UNIQUE` is given, it specifies the number of unique values present for the attribute on any matching object.

It is not an error to specify these conditions on single-valued attributes. However, if you do, the parser warns you that the condition is always true or false.

## Access mode attribute

Specifies an attribute that can be used in a search expression for creating infosets that are based on certain file attributes that are collected during harvests.

Access mode attribute supports these operators: `APPEARS`, `NOT APPEARS`, and `WHERE`.

The `Subject` subattribute supports these operators: `IS`, `IS NOT`, `=`, `!=`, `any()`, and `all()`.

The `Readable` and `Writable` subattributes both support these operators: `IS`, `IS NOT`, `=`, and `!=`.

For example,

- To find data objects for which you collect or maintain information about the read or write file system flags, enter `att: "Access Mode" appears`. This search returns every system-level object on CIFS or NFS volumes.
- To find data objects that are writable by either of the three types of subjects, enter `att: "Access Mode" where (Subject is any (Owner, Group, Others) and Writable is True)` or `att: "Access Mode" where (Writable is False)`. These two examples are semantically identical. If you think of the r/w/x flags as separate groups of flags (that is, `rw` | `rx` | `x`), where one is for the Owner, one is for the Group that a user might be associated with, and one is for Others in that order, then the `Readable` or `Writable` subattributes are evaluated independently for each three group of flags.

The `Subject` subattribute (if present) then just defines which group out of the three must be considered or ignored. The lack of the `Subject` subattribute indicates that the objects must be considered responsive to the expression if the `Readable` or `Writable` conditions (or any combination of them) evaluate to True in ANY of the three groups.

- To find data objects that are read-only by everyone, enter `att: "Access Mode" where (Subject is all (Owner, Group, Others) and Readable is True and Writable is False)`. The subattributes are all ANDed together and the `Subject` subattribute lists every possible option with the `all` qualifier.
- To find data objects that are read-only by at least one type of subject, enter `att: "Access Mode" where (Subject is any (Owner, Group, Others) and Readable is True and Writable is False)` or `att: "Access Mode" where (Readable is True and Writable is False)`. Again, the two examples here produce identical results for the same reason as explained in the previous bullet point.
- To find data objects that are read-only by the owner, enter `att: "Access Mode" where (Subject = Owner and Readable is True and Writable is False)`. You care only about the owner in this case, so you explicitly call `Owner` and only `Owner` out.
- To find data objects that owner can either read or write (not necessarily both), enter `att: "Access Mode" where (Subject = Owner and (Readable is True or Writable is True))`.
- To find data objects that can be written by the owner but is read-only by the group, enter `att: "Access Mode" where ((Subject = Owner and Writable = True) and (Subject = Group and Writable is False))`.



## Type of object to match

For any attribute condition, you can specify what object that condition must match.

You can specify `IN object-type` after the expression. For example, in `att: Extension = doc=in system`, the object must be a doc file at the system level. The condition-scope can be any of these object types here.

Table 7. Available object types	
Available Object Types	Description
all	The matching object can be any type of object. This object is the default.
container	The matching object must be a container object that can have other objects nested inside it. File-archive, email-archive, email, calendar, contact, and task objects are included.
simple	The matching object must not be a container object.
system	The matching object must be a system-level object.
contained	The matching object must be nested within some other container object.

Only one condition-scope can apply to any condition. If you give multiple condition-scopes, the parser issues a warning.

## Multiple matching values for operators

You can give multiple matching values that use the following operators.

- `=`
- `!=`
- `contains`
- `not contains`
- `matches`
- `not matches`
- `starts with`
- `not starts with`
- `ends with`
- `not ends with`

To specify multiple matching values for these operators, give the keyword `any` or `all` after the operator, then a list of values in parentheses. The `any` keyword means that any of the values are a match. The `all` keyword means that all the values must be present. The `all` keyword is only meaningful for multivalued attributes. For example:

Table 8. Keyword examples	
Example	Result
<code>att:"System metadata/Owner" = any (joe, bob, sally)</code>	Owner can be joe, bob, or sally.
<code>att:Size = any (1mb, 2mb, 1gb)</code>	Size can be 1mb, 2mb, or 1gb.

Table 8. Keyword examples (continued)	
Example	Result
att: "Modified date" = any (1 week ago, 2 months ago)	
att: "Accessed date" = any (2009-09-19, 2009-09-20, 2009-09-21)	
att:To = all (joe, bob)	"To" must contain both joe and bob

For any, the operator must be true for at least one of the values. It is as if the user spelled out attribute oper val1 OR attribute oper val2 OR ...

- For all, the operator must be true for all the values. It is as if the user spelled out attribute oper val1 AND attribute oper val2 AND....
- For the NOT various operators, this syntax is the same as using NOT around the entire expression and the corresponding positive operator. For example, att:Size != any( 1mb, 2mb, 2gb ) is the same as NOT (att:Size = any( 1mb, 2mb, 1gb )). The NOT matches all documents that do not match the positive operator.
- It is not an error to specify all on single-valued attributes. However, if you do, the parser warns you that the condition is always false.

## At-least expressions

On occasion, it is helpful to specify that some number of conditions must be true before an object is considered to match. These expressions are called at-least expressions because at least X subexpressions must match an object before the object matches the overall expression.

The syntax for an at-least expression is to give the keywords "at least" followed by a number, and then by a list of expressions in parentheses. For example:

```
at least 2 (att: size=1mb, att: "system metadata/owner"=joe, att:
"object name"=foo.doc, att: "modified date">2009-03-28 )
```

This example specifies that at least two of the four conditions must be true before any object is considered to match the overall expression.

Each subexpression in the list can be arbitrarily complex, but normally they are one condition. Each subexpression in the list must be separated from the others with a comma.

## Parent and child conditions

This parent and child conditions represent advanced function that must be undertaken only by experienced administrators.

You might want to filter for objects whose parents or children match certain conditions. For example, find all emails to Joe in a saved-emails.pst file or find emails to Joe that have an attachment that is named big-contract.pdf. To specify these kinds of queries, give the conditions for the matching objects as normal and add a prefix parent or child to the conditions that apply to the parent or child objects. Here are the expressions for the two previous examples:

```
att:"any address"=Joe and parent att: "object name"="saved-
emails.pst"

att:To=Joe and child att:"object name"="big-contract.pdf"
```

If there is more than one expression that must match the parent or child objects, you must put all the conditions inside parentheses, for example:

```
att:"any address"=Joe and parent(att: "object name"="saved-
emails1.pst" or att:"object name"="saved-emails2.pst")

att:To=Joe and child (att:"object name"="big-contract.pdf" and
att:"modified date">2009-09-07)
```

It is also possible to specify conditions on the object level of parent or child objects. An object level is the nesting level of an object. System-level objects are at level one , immediately nested objects are at level two, and so on. For example, if you want to find the non-nested emails to Joe in saved-emails.pst, you can do:

```
att:"any address"=Joe and att:"object level" = 2
and parent
att:"object name"="saved-emails.pst"
```

It specifies that the emails to Joe must be at object level 2 and be inside saved-email.pst.

It is also possible to specify the object level of the parent or child object relative to the other. For example, to find the parent object of any file that is named nested-file.doc, you can specify it this way:

```
child att: "object name" = "nested-file.doc" parent-level = child-level - 1
```

It means that nested-file.doc can itself be at any level, and the query results are the direct parent of those files.

The general form of relative parent-child level conditions is listed here.

```
parent-level operator child-level +|- number
child-level operator parent-level +|- number
```

```
parent-level operator child-level +|- number
child-level operator parent-level +|- number
```

- Plus (+) in the +|- indicates that the parent-child must be more deeply nested than the other objects.
- Minus (-) means the parent-child must be less deeply nested than the other objects.

The operator can be any of these comparisons:

Numeric Comparisons for Operators					
=	!=	<	<=	>	>=

## No-inheritance conditions

The no-inheritance conditions represent advanced function that must be undertaken only by experienced administrators.

Normally, the search engine allows some objects to inherit the values of their parent objects. Most of the time this behavior is intuitive, but there are times when this behavior must be turned off. When attribute inheritance is not wanted, prefix the expression with `noinheritance`, and attribute inheritance is turned off for that expression. For example, this expression finds emails to Joe containing stock option in the body or any attachment:

```
att: To=joe and "stock option"
```

However, if you want to find emails to Joe with stock option in the email body itself, then you must turn off inheritance for that part of the expression:

```
noinheritance ( att: To=joe and "stock option" )
```

## Sampling

Checking a small, randomly selected subset of an otherwise huge info set for compliance or violation of documents provides quicker results than scanning the whole info set. Create a statistically representative random subset of an info set by modifying any filter with a SAMPLE expression.

**Restriction:** You can use sampling filters only in IBM StoredIQ Data Workbench, not in IBM StoredIQ Insights. However, info sets created with a sampling filter are available in IBM StoredIQ Insights just like any other info sets.

To create a sampling filter, append the SAMPLE keyword together with the percentage of documents that you want to be taken as sample to your query:

```
filter expression SAMPLE percentage
```

### ***filter expression***

Any IBM StoredIQ filter expression but you can choose a condition that always returns true.

### ***percentage***

The percentage of documents to contribute to the sample expressed as float value between 0 and 1.

Filter examples:

```
(att: "/Library/Attributes/System metadata/Size" > 20 KB IN all) SAMPLE 0.1  
(att: "/Library/Attributes/System metadata/Size" >= 0 KB IN all) SAMPLE 0.1
```

If *filter expression* returns 1,000 documents, *filter expression* SAMPLE 0.1 returns 100 documents.

If *filter expression* returns zero documents, *filter expression* SAMPLE 0.1 also returns zero documents.

If *filter expression* returns only a few documents, the percentage might be rounded up: a 0.1 sample of 5 documents will be 1 document because you can't have 0.5 documents.

Such a modified filter returns a randomized subset of the base filter expression results; whenever you apply the filter, a different set of documents is returned. The size of the result set remains the same. The result set is a regular info set on which you can run any IBM StoredIQ action or report.

Sampling from info sets that span multiple volumes is also supported. While the resulting info set will contain a representative random subset with results from each volume, the order of documents within the resulting info set might not be randomized but grouped by volume. When you preview the results of a sampling filter applied to an info set that spans multiple volumes, sampling is done on the final result set from all volumes. In this case, the order of documents within the filter results preview is randomized. Depending on how the sampling filter results are ordered in the final result set, the preview might not have results from all volumes.

## Notices

---

This information was developed for products and services offered in the U.S.A. This material may be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504-1785  
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
US

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

#### COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows:

© (your company name) (year).

Portions of this code are derived from IBM Corp. Sample Programs.

© Copyright IBM Corp. \_enter the year or years\_.

## Trademarks

---

IBM, the IBM logo, and [ibm.com](http://www.ibm.com) are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" <http://www.ibm.com/legal/copytrade.shtml>.

Adobe and PostScript are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

UNIX is a registered trademark of The Open Group in the United States and other countries.

VMware, VMware vCenter Server, and VMware vSphere are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

## Terms and conditions for product documentation

---

Permissions for the use of these publications are granted subject to the following terms and conditions.

### **Applicability**

These terms and conditions are in addition to any terms of use for the IBM website.

### **Personal use**

You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative work of these publications, or any portion thereof, without the express consent of IBM.

### **Commercial use**

You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.

### **Rights**

Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.

IBM reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by IBM, the above instructions are not being properly followed.

You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.

IBM MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.

## IBM Online Privacy Statement

---

IBM Software products, including software as a service solutions, ("Software Offerings") may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering's use of cookies is set forth below.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, See IBM's Privacy Policy at <http://www.ibm.com/privacy> and IBM's Online Privacy Statement at <http://www.ibm.com/privacy/details> the section entitled "Cookies, Web Beacons and Other Technologies" and

the “IBM Software Products and Software-as-a-Service Privacy Statement” at <http://www.ibm.com/software/info/product-privacy>.



---

# Index

## A

- access mode attribute [28](#)
- administration
  - tag groups [6, 7](#)
  - tags [6, 7](#)
- ASCII characters [22](#)
- at-least expressions [30](#)
- attribute [28](#)
- attribute condition [23](#)
- attribute conditions [24](#)

## C

- charts
  - analytics [13](#)
  - basic [12](#)
  - Insights [2, 3, 11–13](#)
- compound attribute [24](#)
- concepts
  - facets [2](#)
  - Insights [2, 3, 11](#)
  - tag groups [2](#)
  - tags [2](#)
- condition operators [24](#)
- conditions
  - compound attributes [27](#)
  - date attributes [25](#)
  - enumeration attributes [25](#)
  - multivalued sparse or dense attributes [27](#)
  - no-inheritance [31](#)
  - number attributes [26](#)
  - parent and child conditions [30](#)
  - sparse attributes [27](#)

## D

- dense attribute [24](#)
- documents
  - tag groups [8](#)
  - tags [8](#)

## E

- expression language [16](#)

## F

- facets
  - Insights [3, 11](#)
- filter
  - full-text syntax [16](#)
- filter condition reference [16](#)
- filter mode
  - Insights [2, 3, 11](#)
- full-text filter [19](#)

- full-text macro [19](#)

## I

- Insights
  - charts [3, 11–13](#)
  - concepts [2, 3, 11](#)
  - evaluating [5](#)
  - facets [2, 3, 11](#)
  - filter mode
    - charts [2, 3](#)
    - labels [2, 3](#)
    - tagging [2, 3](#)
    - tags [2, 3](#)
    - visualization [2, 3](#)
  - results [2, 3, 11–13](#)
  - reviewers [2](#)
  - search [2](#)
  - search options [3](#)
  - searching [5](#)
  - tagging [6](#)
  - visualization [3, 11–13](#)

## L

- labels
  - Insights [2](#)
- legal
  - notices [33](#)

## M

- macro [19](#)
- malformed query [6](#)
- multiple matching values [29](#)
- multivalued attribute [24](#)

## N

- no-inheritance conditions [31](#)
- notices
  - legal [33](#)

## O

- operators [29](#)

## P

- parent and child conditions [30](#)

## Q

- queries
  - regex [21](#)

## R

- re: queries [21](#)
- reference [23](#)
- regex [21](#)
- regular expressions [21](#)
- results
  - Insights [2](#), [3](#), [11](#)
- reviewer
  - tag groups [7](#)
  - tags [7](#)
- reviewers
  - tag groups [10](#)
  - tags [10](#)

## S

- search
  - Insights [2](#), [3](#)
- single-valued attribute [24](#)
- sparse attribute [24](#)
- specification of attribute conditions [24](#)
- string attributes [24](#)
- syntax error [6](#)

## T

- tag groups [2](#)
- tagging
  - documents [8](#)
  - Insights [2](#)
  - reviewers [10](#)
- tags
  - administration [6](#), [7](#)
  - Insights [2](#)
  - tag groups [6](#), [7](#)

## V

- visualization
  - Insights [2](#), [3](#), [11–13](#)



