Conceptual analytics setup basics

This quick reference guide contains basic workflows for setting up an Analytics conceptual index and the conceptual analytics tools. For more detailed information, see the Analytics section of the documentation site.

Analytics index setup

The analytics index is the cornerstone for all the conceptual analytics functionality available in Relativity.

Creating a data source

Set up a saved search which will become the data source for the index. This searches for documents with Extracted Text smaller than 30 MB.

Use the following conditions and fields:

- **Name**—enter a name for your search such as Conceptual Analytics Index Search.
- **Conditions**
  - Extracted Text Size > 0
  - Extracted Text Size < 30000
- **Fields**—Extracted Text

> **Note:** Including the Extracted Text Size field assumes you used Relativity Processing. If you did not use Relativity Processing, you need to populate the Extracted Text Size field with data.

Creating the Analytics index

Use the following settings to set up the Analytics index:

- **Index Information**
  - **Name**—enter Conceptual Analytics Index.
  - **Index Type**—select **Conceptual**.
  - **Data source**—select the saved search created above.
  - **Order**—set the order or leave as default.
- **Advanced Settings**
  - **Training data source**—this will default to the same saved search as the data source.
  - **Optimize training set**—set to **Yes**.
  - **Dimensions**—leave as default.
  - **Remove English signatures and footers**—set to **Yes**.
  - **Enable email header filter**—set to **Yes**. Note that you can't alter this setting unless you set the previous setting to No.
  - **Stop words**—leave as default.
- **Optional Settings**
  - **Email notification recipients**—enter any email address for those who need to be contacted when the index has been created or if there is an error.

Adding repeated content filters

Repeated content filters are created in structured analytics and can also be added by hand. They are added to an Analytics index to remove any boilerplate text that would distort the conceptual relationship of the documents. Please see the Relativity documentation site for information on running automatic repeated content identification in structured analytics.

To add a repeated content filter to an existing index:

1. Click on the **Repeated Content Filters** tab in the bottom panel of the console.
2. Click **Link**.
3. On the Repeated Content Filter modal, find and select the repeated content filters to link to the profile. We recommend sorting by largest number of occurrences.
4. Click **Apply**.

Analytics categorization sets

Analytics categorization uses manually selected example documents as a basis for identifying and grouping other conceptually similar documents. Unlike clustering, you can use categorization to place documents into multiple categories if a document is a conceptual match with more than one category. It also takes user decisions on a few example documents and categorizes all other documents based on those decisions.

Creating workspace fields

Unless your template has the fields listed below, you must create them to use categorization sets in a project.

Field name	Object	Field type
Category designation	Document	Single choice or multiple choice
Category indicator	Document	Yes/No

Additionally, you must add choices for the **Category Designation** field.

Creating a category designations layout and assigning examples

Have your reviewers go through a sample of documents and assign designations to those determined to be good examples for the category. Tag at least 5 to 20 documents and no more than a few thousand. For explanation about what makes a good example document, see Analytics categorization sets on the documentation site.

Note: Only documents that are part of the data source of your conceptual index can serve as examples.

If you want to add a pre-existing Issue field to a categorization set, you may want to conduct a QC review of the example documents that will be used for categorization. We recommend setting up a designated layout for reviewers to tag documents, and do not use the right-click functionality in the viewer.

Sample layout

- **Layout name**—enter a name for the layout.
- **Fields**
  - Category Designation
  - Category Indicator

Creating the categorization set

The Analytic categorization set is used to identify the documents, Analytics index, fields and other criteria needed for categorization

Use the following settings to set up the Analytics categorization set:

- **Name**—enter a name for the categorization set.
- **Documents To Be Categorized**—select the saved search containing the documents to be categorized. Only include documents that are also part of the data source of your Analytics index.
- **Analytics Index**—select the index you created that defines the conceptual space.
- **Minimum Coherence Score**—leave as default.
- **Maximum Categories Per Document**—leave as default.
- **Email notification recipients**—enter any email address for those who need to be contacted when categorization is complete.
- **Categories and Examples Source**—select the Category Designation field.
- **Example Indicator Field**—select the Category Indicator field.

Clustering

Once the Analytics index has been created, you can execute the clustering mass action from the Document list. You can create multiple clustering sets based on a search or sub-set of documents.

To use the clustering mass action:

1. Navigate to the Documents tab and find the documents you want to cluster.
2. Select the documents you want to cluster individually or select **All** from the drop-down menu.
3. Select **Cluster** from the mass operations drop-down menu.
4. Select **Cluster**.
5. Complete the following fields:
    - **Cluster Options**
        - **Name**—enter a name for the cluster set
        - **Relativity Analytics Index**—select the conceptual index that includes the documents you want to cluster.
    - **Advanced Options**
        - **Title Format**—leave as default.
        - **Maximum Hierarchy Depth**—leave as default.
        - **Minimum Coherence**—leave as default.
        - **Generality**—leave as default.
        - **Create Cluster Score Field**—leave this as **No** unless you want to know each document's coherence score. Setting this to Yes significantly slows down the clustering process.
6. Click **Submit for Clustering**.

Find similar documents, keyword expansion, and concept searching

These conceptual analytics tools do not require any additional setup. For more information on them, see the Analytics section of the documentation site.

Proprietary Rights

This documentation ("**Documentation**") and the software to which it relates ("**Software**") belongs to Relativity ODA LLC and/or Relativity's third party software vendors. Relativity grants written license agreements which contain restrictions. All parties accessing the Documentation or Software must: respect proprietary rights of Relativity and third parties; comply with your organization's license agreement, including but not limited to license restrictions on use, copying, modifications, reverse engineering, and derivative products; and refrain from any misuse or misappropriation of this Documentation or Software in whole or in part. The Software and Documentation is protected by the **Copyright Act of 1976**, as amended, and the Software code is protected by the **Illinois Trade Secrets Act**. Violations can involve substantial civil liabilities, exemplary damages, and criminal penalties, including fines and possible imprisonment.

©2026. Relativity ODA LLC. All rights reserved. Relativity® is a registered trademark of Relativity ODA LLC.