

Analytics Index Quick Reference Guide

This guide provides step-by-step instructions for creating an Analytics conceptual index in order to run conceptual analytics tools (concept searching, find similar documents, categorization, clustering, and keyword expansion).

In addition, this guide explains the different stages of index creation.

Creating an Analytics conceptual index

1. Create a saved search to serve as your searchable set. The searchable set is the collection of documents on which you want to perform any conceptual analytics operation. This search should only pull back authored content fields (ex. extracted text) and typically needs no additional conditions.
2. Click the **Indexing & Analytics** tab and select **Analytics Indexes**.
3. Click **New Analytics Index**. The Analytics Index Information form appears.
4. Complete the following fields on the Analytics Index Information form.
 - a. **Name** – enter a name for the index.
 - b. **Order** – enter an order for the index. The order determines the relative position of the index in the search drop-down along with other search providers such as dtSearch and keyword search.
 - c. **Index type** – select **Conceptual**.
 - d. **Searchable set** – select the saved search you created in Step 1. The Training set field automatically populates with this same saved search.

Leave all other fields under Advanced Settings as default.

***Note:** For a complete description of the Analytics index fields, see the Relativity documentation site.*

5. Click **Save**.
6. Click **Populate Index: Full**.

Index stages

Index creation consists of three stages: Population, Building, and Activating.

Note: By default, Analytics indexes move through all three stages automatically. If you want to disable this behavior, modify the index settings **Continue index steps to completion**. For more information, see the *Relativity documentation site*.

The following steps occur during each stage of index creation:

Population

- All documents from the training set and searchable set are staged and flagged for pre-processing.
- Document pre-processing occurs to clean up text in the following ways:
 - Numbers and symbols are ignored.
 - All words are made lowercase.
 - Filters found under Advanced Settings are applied (ex. email header filter).
 - Repeated content filters are applied.

Building

- Training set documents and Latent Semantic Indexing (LSI) are used to build the concept space based on the relationships between words and documents.
- Searchable set documents are mapped into the concept space.
- Concept stop words (very common words) are filtered from the index to improve quality.

Activating

- Makes the index active and available in the search indexes drop-down.
- Saves the index to RAM which loads the index into memory. If you find yourself running out of free RAM on the Analytics server, deactivate your index.
 - Analytics indexes are automatically deactivated after 30 days of inactivity. You can reactivate the index from the index console.

Common workflows

There may be times when you need to update your index. Depending on the update you're making, you can save time by running an incremental population or only running a build. The following table outlines various workflows for different index updates.

Workflow	Index update
Adding new documents that: <ul style="list-style-type: none">• Introduce new concepts• Make up more than 10% - 30% of your document population	<ol style="list-style-type: none">1. Add documents to both the training and searchable set saved search.2. Click Populate Index: Incremental.
Adding new documents that: <ul style="list-style-type: none">• Don't introduce new concepts• Make up less than 10% - 30% of your document population	<ol style="list-style-type: none">1. Add documents to the searchable set saved search only.2. Click Populate Index: Incremental.
Removing documents from the training or searchable set	<ol style="list-style-type: none">1. Remove documents from the training or searchable set saved search.2. Click Populate Index: Incremental.
Updating concept stop words	<ol style="list-style-type: none">1. Click Deactivate Index.2. Click Build Index.
Updating extracted text (ex. Updating poor quality OCR text)	<ol style="list-style-type: none">1. Update extracted text.2. Click Populate Index: Full.
Updating filters (email header, repeated content)	<ol style="list-style-type: none">1. Update filters.2. Click Populate Index: Full.