coveo™

Coveo Platform 7.0

YouTube Connector Guide

Notice

The content in this document represents the current view of Coveo as of the date of publication. Because Coveo continually responds to changing market conditions, information in this document is subject to change without notice. For the latest documentation, visit our website at www.coveo.com.

© Coveo Solutions Inc., 2015

Coveo is a trademark of Coveo Solutions Inc. This document is protected by intellectual property laws and is subject to all restrictions specified in the Coveo Customer Agreement.

Document part number: PM-150216-EN
Publication date: 1/3/2019

Table of Contents

1. YouTube Connector

CES 7.0.7433+ (February 2015)

The Coveo connector for YouTube allows Coveo administrators to index and integrate YouTube user and brand channels content into the Coveo unified index. The connector indexes all items from specified YouTube channel(s) so that in the Coveo search interfaces, a user can easily find this content.

1.1 Connector Features Summary

Features		Supported	Additional information
YouTube version		Data API (v3)	
Searchable content types		✔	Metadata on videos, channel playlists, and playlist items.
Content update	Incremental refresh	✔	Full refresh or rebuild needed to retrieve: • Deleted videos • Changes on videos • Statistics on video (e.g., view counts, likes, etc.) updates • Playlist and playlist item updates.
	Full refresh	✔	
	Rebuild	✔	
Document-level security		✖	Permissions must be manually defined on the source [more]

1.2 Features

The YouTube connector features are:

Content indexing

- Extraction and indexing of all available YouTube metadata on videos (including thumbnails).

- CES 7.0.7711+ (June 2015) YouTube channel playlists and playlist items are indexed, but channel related playlists, such as the **Likes** and **Watch Later**, are not indexed.

Partial Incremental refresh

YouTube channels are periodically re-indexed by the connector to catch newly uploaded videos (see "Configuring and Indexing a YouTube Source" on page 9).

Notes:

- Private and unpublished videos are not indexed.

- Deleted and changed videos require a full refresh to be taken in account (see Limitation).

Pause/Resume

When indexing YouTube content, the connector can be paused and resumed.

1.3 Limitation

- Limited incremental refresh capabilities:

  - A full refresh is needed to update deleted videos

  - A full refresh is needed to update changes on videos

  - A full refresh is needed to update statistics on videos (e.g., view counts, likes, ...)

  - A full refresh is needed to update playlists and playlist items.

    Note: CES 7.0.7711+ (June 2015) Playlist and Playlist items of a YouTube channel are indexed.

What's Next?

Review the steps to deploy the YouTube connector (see "YouTube Connector Deployment Overview" on page 3).

2. YouTube Connector Deployment Overview

The following procedure outlines the steps needed to deploy the YouTube connector. The steps indicate the order in which you must perform configuration tasks on both the Google and Coveo servers.

To deploy the YouTube connector

1. Validate that your environment meets the requirements (see "YouTube Connector Requirements" on page 4).

2. On the Google server, create a Google API Console project to allow the Coveo connector access to the YouTube channel(s) of your choice (see "Allowing the Coveo Connector Access to YouTube Data" on page 5).

3. On the Coveo server, in the Coveo Administration Tool:

   a. Create a YouTube field set to take advantage of the available YouTube metadata.

      i. It is recommended to start by importing the default YouTube field set file (`[CES_Path]\Bin\Coveo.CES.CustomCrawlers.YouTube.FieldSet.xml`) to create fields for all the metadata available by default from YouTube videos.

      ii. When you created custom metadata for your YouTube documents, add corresponding fields to the field set.

   b. Configure and index a YouTube source.

      The connector must know details about the YouTube channels to index their content (see "Configuring and Indexing a YouTube Source" on page 9).

   c. If you encounter issues, verify if modifying the default value of hidden source parameters can help resolve the problems (see "Modifying Hidden YouTube Source Parameters" on page 14).

3. YouTube Connector Requirements

Your environment must meet the following requirements to be able to use the YouTube connector:

- CES 7.0.7433+ (February 2015)

- Coveo license for the YouTube connector

  Your Coveo license must include support for the YouTube connector to be able to use this connector.

- A valid Google Account

  Using a Google Account, you must log in the Google API Console to authorize Coveo access to YouTube data (see Allowing the Coveo Connector Access to YouTube Data).

What's Next?

The YouTube connector needs an API (YouTube Data API v3) to identity itself to Google. The connector uses this API, which is a link between the connector and the YouTube channels, to access the public videos that you want to index (see "Allowing the Coveo Connector Access to YouTube Data" on page 5).

4. Allowing the Coveo Connector Access to YouTube Data

The Coveo connector uses a Google API (YouTube Data API v3) to retrieve information about the different items in YouTube (videos, users, channels). You must enable the YouTube Data API v3 API and create an API key to authorize the Coveo connector to access the YouTube public content that you want to index.

Note: The YouTube Data API v3 API does not provide account information or give any sort of authentication rights.

To allow the connector access to YouTube data

1. Go to the Google Developers Console, and log in using a Google Account.
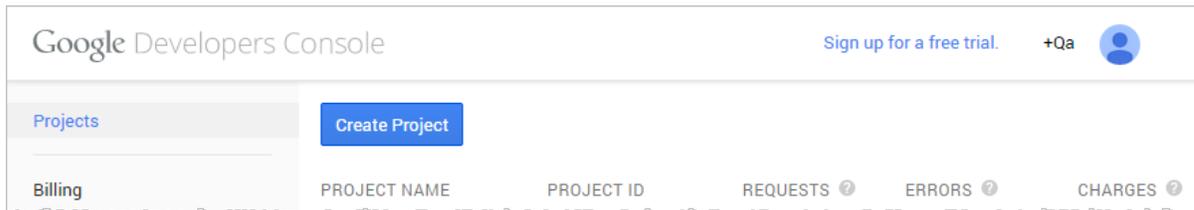
2. Create a service account project for the Coveo connector:

   Note: You can also select an existing project and start this step from substep c.

   a. In the **Projects** page, click **Create Project**.

   b. In the **New Project** dialog box, enter a project name and ID, and then click **Create**.
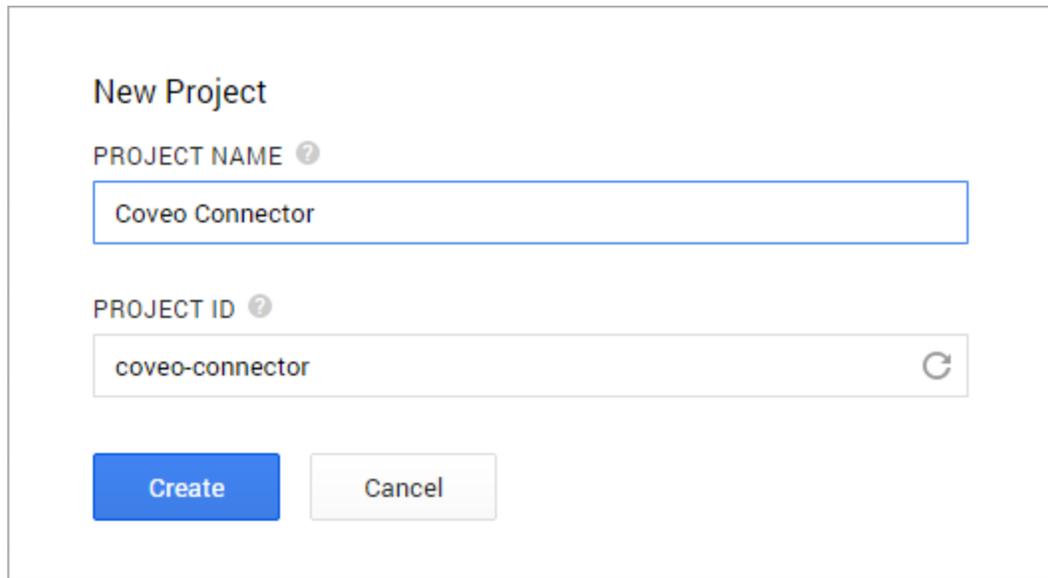
   c. Click your **[Project Name]**.

   d. Access the **APIs** page:

- In the sidebar on the left, click **APIs & auth**, and then click **APIs**.

  OR

- In the **Project Dashboard** page, click **Enable an API**.

e. In the **APIs** page, ensure that for **YouTube Data API v3**, the **Status** is set to **ON**.
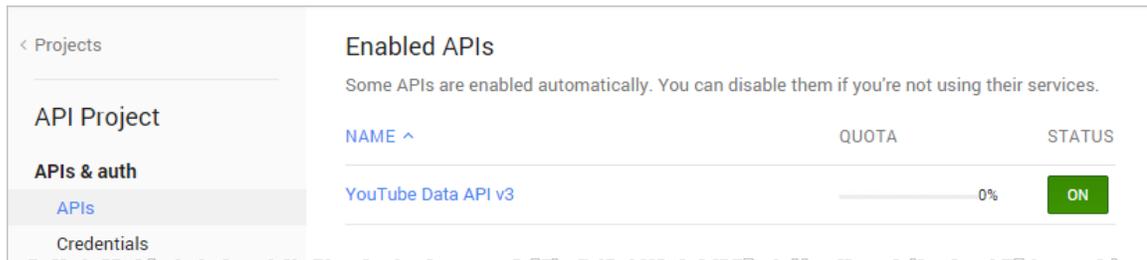
3. In the sidebar on the left, under **APIs & auth**, select **Credentials**, and then in the **Public API access** section, click the **Create new key** button.
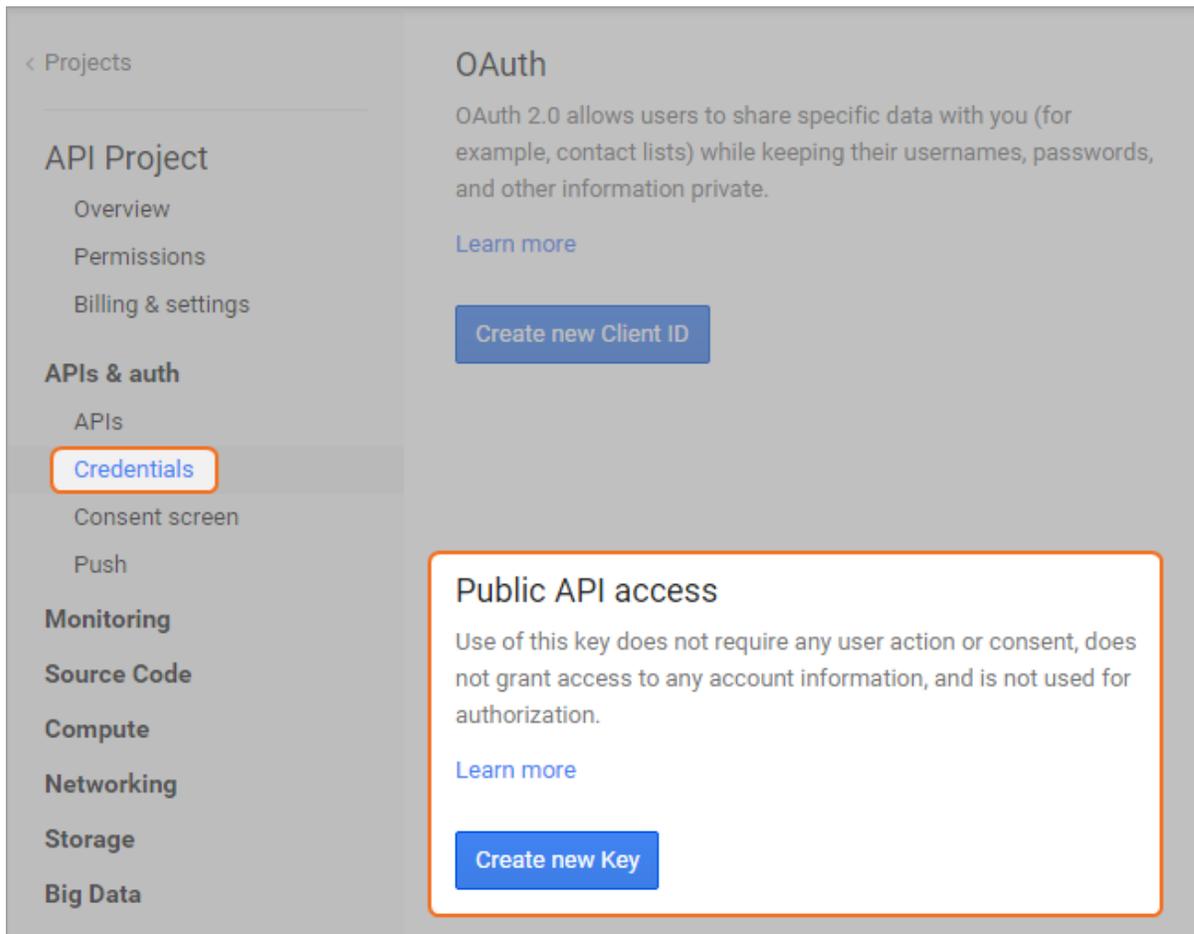
4. In the next **Create a new key** dialog box, click the **Server key** button.
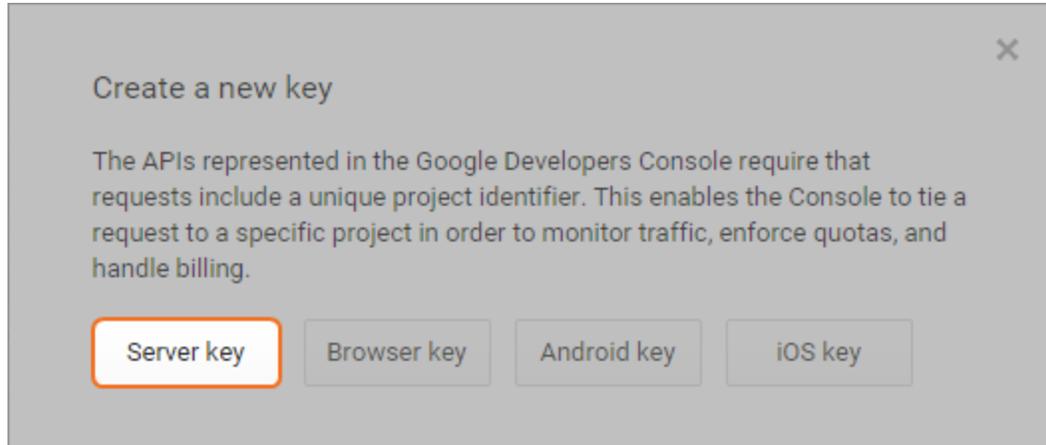
5. In the **Create a sever key and configure allowed IPs** dialog box that appears, click **Create**.

Note: When you want to limit the utilization of your API key, you can enter your server path, meaning that only the requests from your server will be accepted.
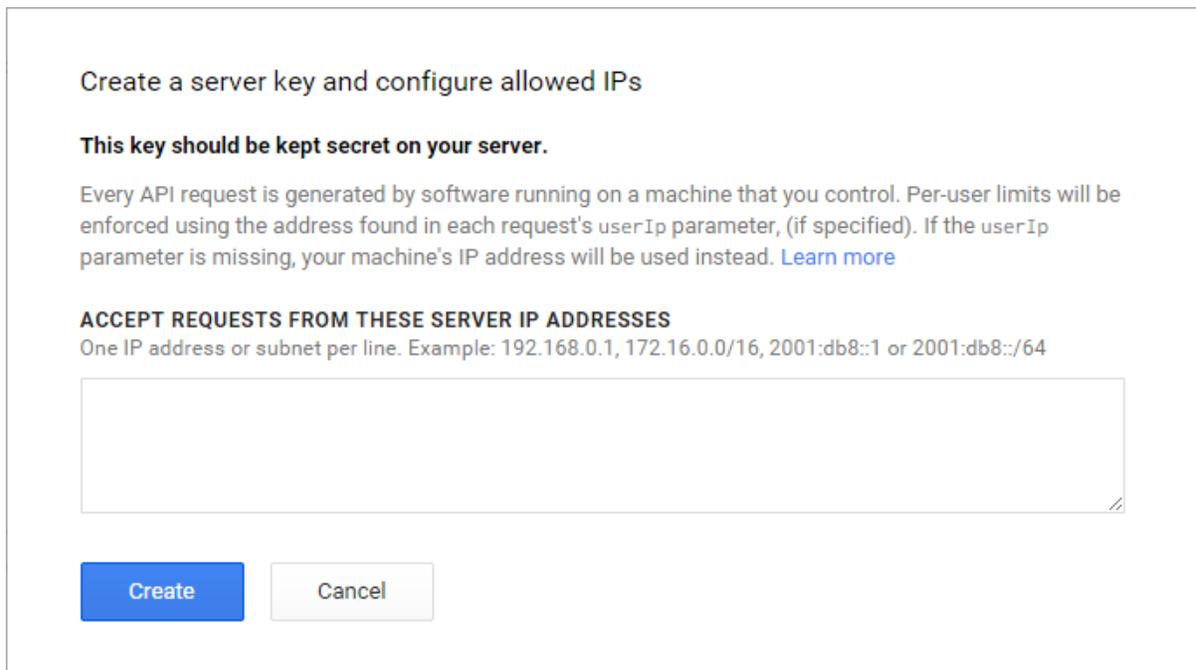
6. Back in the **Google Developers Console**, in the **Key for server applications** section, take note of the **API key** that you will need later to configure your YouTube source.

Key for server applications

API KEY	
IPS	Any IP allowed
ACTIVATION DATE	Feb 12, 2015, 6:12:00 AM
ACTIVATED BY	

Edit allowed IPs   Regenerate key   Delete

5. Configuring and Indexing a YouTube Source

A source defines a set of configuration parameters for a specific YouTube channel.

To configure and index a YouTube source

1. On the Coveo server, access the Administration Tool.

2. Select **Index** > **Sources and Collections**.

3. In the **Collections** section:

    Select an existing collection in which you want to add the new source.

    OR

    Click **Add** to create a new collection.

4. In the **Sources** section, click **Add**.

    The **Add Source** page that appears is organized in three sections.

5. In the **General Settings** section of the **Add Source** page:

a. Enter the appropriate value for the following required parameters:

    Name

        Enter a descriptive name of your choice for the connector source.

Example: `The Company YouTube Channel`

Source Type

Select the connector used by this source. In this case, select **YouTube**.

Note: If you do not see **YouTube**, your environment does not meet the requirements (see "YouTube Connector Requirements" on page 4).

Addresses

Enter the URL of one or more YouTube user or brand channels, or playlists.

Example:

- `https://youtube.com/channel/channelID`

- `https://youtube.com/user/username`

- `https://youtube.com/playlist?list=playlistID`

Notes:

- When you enter a URL pointing to a specific user who manages more than one channels, the connector indexes all videos under these channels.

- CES 7.0.8541+ (September 2016) The connector supports playlist URLs.

Fields

Select the field set that you created for your YouTube source (see YouTube Connector Deployment Overview).

Refresh Schedule

Time interval at which the index is automatically refreshed to keep the index content up-to-date. By default, the **Every day** option instructs CES to refresh the source everyday at 12 AM. Because the incremental refresh takes care of maintaining the source up-to-date, you can select a longer interval such as **Every Sunday**.

b. Review the value for the following parameters that often do not need to be modified:

Rating

Change this value only when you want to globally change the rating associated with all items in this source relative to the rating to other sources.

Example: When a source replaces a legacy system, you may want to set this parameter to **High**, so that in the search interface, results from this source appear earlier in the list compared to those from legacy system sources.

Document Types

If you defined a custom document type set for this source, select it.

Active Languages

If you defined custom active language sets, ensure to select the most appropriate for this source.

6. In the **Specific Connector Parameters & Options** section of the **Add Source** page:

a. In the **API Key** box, enter the API key you previously created (see Allowing the Coveo Connector Access to YouTube Data).

b. In the **Mapping File** box, the path to the default mapping file that defines how the connector handles metadata often does not need to be changed.

c. CES 7.0.8388+ (June 2016) Select the **Index Playlists** check box when you want YouTube channel playlists and playlist items to be indexed.

> **Notes:**
>
> - CES 7.0.8541+ (September 2016) This option is not used when you specify a playlist URL in the **User, Channel, or Playlist URL** parameter.
>
> - CES 7.0.8225– (March 2016) YouTube channel playlists and playlist items are indexed by default.

d. Click **Add Parameter** when you want to show and change the value of advanced source parameters (see "Modifying Hidden YouTube Source Parameters" on page 14).

e. The **Option** check boxes generally do not need to be changed:

Index Subfolders CES 7.0.7599– (April 2015)

This parameter is not taken into account for this connector.

Index the document's metadata

When selected, CES indexes all the document metadata, even metadata that are not associated with a field. The orphan metadata are added to the body of the document so that they can be searched using free text queries.

When cleared (default), only the values of system and custom fields that have the **Free Text Queries** attribute selected will be searchable without using a field query.

> **Example:** A document has two metadata:
>
> - `LastEditedBy` containing the value `Hector Smith`
>
> - `Department` containing the value `RH`
>
> In CES, the custom field `CorpDepartment` is bound to the metadata `Department` and its **Free Text Queries** attribute is selected.
>
> When the **Index the document's metadata** option is cleared, searching for `RH` returns the document because a field is indexing this value. Searching for `hector` does not return the document because no field is indexing this value.
>
> When the **Index the document's metadata** option is selected, searching for `hector` also returns the document because CES indexed orphan metadata.

Document's addresses are case-sensitive <mark>CES 7.0.7433– (February 2015)</mark>

Ensure that this option is selected since YouTube video IDs are case sensitive.

Generate a cached HTML version of indexed documents

When you select this check box (recommended), at indexing time, CES creates HTML versions of indexed documents. In the search interfaces, users can then more rapidly review the content by clicking the **Quick View** link rather than opening the original document with the original application. Consider clearing this check box only when you do not want to use **Quick View** links or to save resources when building the source.

Open results with cached version

Leave this check box cleared (recommended) so that in the search interfaces, the main search result link opens the original document with the original application. Consider selecting this check box only when you do not want users to be able to open the original document but only see the HTML version of the document as a **Quick View**. In this case, you must also select **Generate a cached HTML version of indexed documents**.

   f. Click **Save** to save the source configuration.

7. Since the YouTube connector only indexes public videos, security permissions on the source content must be disabled:

a. In the navigation menu on the left, select **Permissions**.

b. Next to **Permissions**, select the **Specifies the security permissions to index** option.

c. Next to **Allowed Users**, ensure that a well-known everyone group such as the Active Directory `everyone \S-1-1-0\` is added.

d. Click **Apply Changes**.

e. On the toolbar, click **Rebuild** to start indexing your source.

8. Validate that the source building process is executed without errors:

   - In the navigation panel on the left, click **Status**, and then validate that the indexing proceeds without errors.

     OR

   - Open the CES Console to monitor the source building activities.

What's Next?

Set an incremental refresh schedule for your source.

Consider modifying some hidden source parameters to try resolving other issues (see "Modifying Hidden YouTube Source Parameters" on page 14).

5.1 Modifying Hidden YouTube Source Parameters

The **Add Source** and **Source: ... General** pages of the Administration Tool present the parameters with which you can configure the connector for most YouTube setups. More advanced and more rarely used parameters are hidden. You can choose to make one or more of these parameters appear in the **Add Source** and **Source: ... General** pages of the Administration Tool so that you can change their default value. Consider changing values of hidden parameters when you encounter issues.

The following list describes the advanced hidden parameters available with YouTube sources. The parameter type (integer, string, etc.) appears between parentheses following the parameter name.

IndexPlaylistItems (Boolean) `CES 7.0.8541+ (September 2016)`

> Whether playlist items are indexed. The default value is `true`. When set to `false`, playlist item titles are only displayed in the `ytplaylistitemstitle` field.

BatchSize (Integer)

> Number of items to fetch per request made to the YouTube server. The default value is `30` items. The minimum value is `1` item. A small value forces the connector to make small but frequent queries to YouTube. A larger value leads to larger and less frequent queries.

LastRefreshCushion (Integer)

> The time set in minutes to ensure the incremental refresh catches all newly uploaded videos. The default value is `180` minutes.

> **Note:** Since the Google API (YouTube Data API v3) has a latency, add this parameter to set the incremental refresh to look for newly uploaded videos during a wider time interval. Without this parameter, if an incremental refresh runs before an uploaded video is available to be caught (YouTube must finish processing the video), the video would not be indexed.

To modify hidden YouTube source parameters

1. Refer to "Adding an Explicit Connector Parameter" on page 15 to add one or more YouTube source parameters.

2. For a new YouTube source, access the **Add Source** page of the Administration Tool to modify the value of the newly added advanced parameter:

    a. Select **Index** > **Sources and Collections**.

    b. Under **Collections**, select the collection in which you want to add the source.

    c. Under **Sources**, click **Add**.

    d. In the **Add Source** page, edit the newly added advanced parameter value.

3. For an existing YouTube source, access the **Source: ... General** page of the Administration Tool to modify the

value of the newly added advanced parameter:

    a. Select **Index** > **Sources and Collections**.

    b. Under **Collections**, select the collection containing the source you want to modify.

    c. Under **Sources**, click the existing YouTube source in which you want to modify the newly added advanced parameter.

    d. In the **Source: ... General** page, edit the newly added advanced parameter value.

4. Rebuild your YouTube source to apply the changes to the parameters.

5.2 Adding an Explicit Connector Parameter

Connector parameters applying to all sources indexed using this connector are called explicit parameters.

When you create or configure a source, the Coveo Enterprise Search (CES) 7.0 Administration Tool presents parameters with which you can configure the connector for most setups. For many connectors, more advanced and more rarely used parameters also exist but are hidden by default. CES then uses the default value associated with each of these hidden parameters.

You can however choose to make one or more of these parameters appear in the **Add Source** and **Source: ... General** pages of the Administration Tool so that you can change their default value.

To add an explicit connector parameter

1. On the Coveo server, access the Administration Tool.

2. Select **Configuration** > **Connectors**.

3. In the list on the **Connectors** page, select the connector for which you want to show advanced hidden parameters.

4. In the **Parameters** section of the selected connector page, click **Add Parameter** for each hidden parameter that you want to modify.

   Note: The **Add Parameter** button is present only when hidden parameters are available for the selected connector.

5. In the **Modify the parameters of the connector** page:

a. In the **Type** list, select the parameter type as specified in the parameter description.

b. In the **Name** box, type the parameter name exactly as it appears in the parameter description. Parameter names are case sensitive.

c. In the **Default Value** box, enter the default value specified in the parameter description.

> **Important:** Do not set the value that you want to use for a specific source. The value that you enter here will be used for all sources defined using this connector so it must be set to the recommended default value. You will be able to change the value for each source later, in the **Add Source** and **Source: ... General** pages of the Administration Tool.

d. In the **Label** box, enter the label that you want to see for this parameter.

> **Example:** To easily link the label to the hidden parameter, you can simply use the parameter name, and if applicable, insert spaces between concatenated words. For the **BatchSize** hidden parameter, enter `Batch Size` for the label.

> **Note:** To create multilingual labels and quick help messages, use the following syntax: `<@ln>text</@>`, where *ln* is replaced by the language initials—the languages of the Administration Tool are English (en) and French (fr).

> **Example:** `<@fr>Chemin d'accès du fichier de configuration</@><@en>Configuration File Path</@>` is a label which is displayed differently in the French and English versions of the Administration Tool.

> **Tip:** The language of the Administration Tool can be modified by pressing the following key combination: `Ctrl+Alt+Page Up`.

e. Optionally, in **Quick Help**, enter the help text that you want to see for this parameter when clicking the question mark button ? that will appear beside the parameter value.

> **Tip:** Copy and paste key elements of the parameter description.

f. When **Predefined values** is selected in the **Type** parameter, in the **Value** box that appears, enter the parameter values that you want to see available in the drop-down parameter that will appear in the Administration Tool interface. Enter one value per line. The entered values must exactly match the values listed in the hidden parameter description.

g. Select the **Optional parameter** check box when you want to identify this parameter as an optional parameter. When cleared, CES does not allow you to save changes when the parameter is empty. This parameter does not appear for **Boolean** and **Predefined values** parameter types.

h. Select the **Sensitive information** check box for password or other sensitive parameter so that, in the Administration Tool pages where the parameter appears, the typed characters appear as dots to mask them. This parameter appears only for the **String** type.

> **Example:** When you select the **Sensitive information** check box for a parameter, the characters typed appear as follows in the text box:
>
> ••••

i. Select the **Validate as an email address** check box when you want CES to validate that the text string that a user enters in this parameter respects the format of a valid email address. This parameter appears only for the **String** type.

j. In the **Maximum length** box, enter the maximum number of characters for the string. This parameter appears only for the **String** type. When you enter `0`, the length of the string is not limited.

k. Click **Save**.

6. Back in the **Connector** page, click **Apply Changes**.

The hidden parameter now appears in the **Add Source** and **Source: ... General** pages of the Administration Tool for the selected source. You can change the parameter value from these pages. Refer to the documentation for each connector for details.

> **Note:** When you want to modify a hidden source parameter, you must first delete it, and then redefine it with the modified values.