Configuring Pulse Video Analytics (Licensed Feature)

Activation To use this feature, you must purchase and install the Pulse video analytics feature license on the Cisco MXE 3500 appliance.

The Pulse video analytics feature enables keyword tagging and speaker identification in video content. The enhanced video content is shared through Cisco Show and Share.

Note An engagement with Cisco Advance Services is required for implementing Pulse video analytics.

Note Integration with Cisco Show and Share Release 5.3.0 is required to use this feature.

This section contains the following information:

Keyword Tagging

Keyword tagging automatically identifies the relevant keywords spoken in a video and tags their location.

The Cisco MXE 3500 uses a vocabulary list to automatically identify and tag spoken words in a video with the matching keywords. The keywords are displayed in Cisco Show and Share, and the viewer can search or go to the specific section of the video where the keyword was spoken by clicking on the keyword.

Workflow for Implementing Keyword Tagging

1. Create a domain specific document repository. This is required to generate a vocabulary list. Use one or both of the following methods to create the document repository:

   – Manually identify documents and create a document repository.

   – Use the provided Crawler tool to search specified web domains and retrieve all stored documents, and copy them to a local document repository.

Tip Use the Crawler tool to discover more documents representing the domain if you cannot manually provide a sufficient number or breadth of documents across relevant topics. Depending on the size of the web domains, it may take several hours for the crawler to create a document repository.

2. Generate a vocabulary list. Use the AutoVocab tool to extract text from the documents and filter the sentences to produce a vocabulary list.

3. Configure the speech recognition engine used by Pulse video analytics with the new vocabulary words. This step is necessary to customize the speech recognition engine to identify and tag speech in a video with the new keywords in the vocabulary list.

Note An engagement with Cisco Advanced Services team is required to update the speech recognition engine used by Pulse. Please contact your Cisco representative for details.

Guidelines and Limitations

- The Crawler and AutoVocab tools retrieve and extract the vocabulary list from documents in text, PDF, PowerPoint presentation, or Microsoft Word formats only.

- A minimum of 1000 documents (approximately 10 million words) that are relevant to the domain are required to create a good vocabulary list.

Create a Document Repository

Create a document repository using either one or both of the following methods:

- Create a Document Repository Manually, page 4-3
- Create a Document Repository Using the Crawler Tool, page 4-3

Create a Document Repository Manually

Before You Begin
Identify the relevant documents to create a vocabulary list.

Copy the Documents

Step 1   SSH to *mxe_IP_address*, and login as **admin**.

Step 2   Select **System Command Prompt** and select **OK**.

Step 3   Enter **su -**.

Step 4   Copy the relevant documents to **/mnt/shared/crawler_output**.

Note   Provide read access to the directories and files in **/mnt/shared/crawler_output**.

Step 5   To run the Crawler tool, go to Create a Document Repository Using the Crawler Tool, page 4-3.

OR

Go to Generate the Vocabulary List Using the AutoVocab Tool, page 4-4.

Create a Document Repository Using the Crawler Tool

Before You Begin
Identify the web domains for the crawler tool to extract documents from.

Limitations
Only enter URLs starting with http://. The crawl fails when secure web domains (web domains with https://) are entered.

Create a local document repository

Step 1   SSH to *mxe_IP_address*, and login as **admin**.

Step 2   Select **System Command Prompt** and select **OK**.

Step 3   Enter **su -**.

Step 4   Enter **cd /mnt/shared/**

Step 5   Enter **sh /opt/cisco/autovocab/scripts/runCrawler.sh /mnt/shared/crawler_output** *URL*

The *URL* is the web domain, for example http://www.cisco.com, the Crawler tool crawls and extracts the documents from.

When completed, a document repository with a directory structure similar to the directory structure of the web domain is created.

Note   Do not change the directory structure of the document repository. Maintaining the same directory structure speeds up subsequent crawls.

If you have multiple web domains, enter each URL separated by a space. For example,
/mnt/shared/crawler_output *URL1 URL2 URL3*

> **Tip**   Do not delete the files in **crawler_output**. This allows for incremental updates and speeds up subsequent crawls.

> **Tip**   Depending on the size of the web domain a crawl can take several hours.

Step 6   Go to Generate the Vocabulary List Using the AutoVocab Tool, page 4-4.

Generate the Vocabulary List Using the AutoVocab Tool

Step 1   SSH to *mxe_IP_address*, and login as **admin**.

Step 2   Select **System Command Prompt** and select **OK**.

Step 3   Enter **su -**.

Step 4   Enter **cd /mnt/shared/**

Step 5   Enter **sh /opt/cisco/autovocab/scripts/runAutoVocab.sh /mnt/shared/crawler_output /mnt/shared/autovocab_output**.

/mnt/shared/crawler_output is the directory containing the source documents for extracting the vocabulary words. The AutoVocab tool generates the vocabulary list, AVTerms.csv, and saves it to **/mnt/shared/autovocab_output**.

> **Tip**   Use Microsoft Excel to add, remove, and edit the vocabulary list.

Step 6   Go to Upload the Vocabulary List, page 4-4.

Upload the Vocabulary List

Step 1   Go to **http://***mxe_IP_address***/pulse/**, where *mxe_IP_address* is the hostname or IP address for the Cisco MXE 3500.

Step 2   Login as **ncpadmin.** Enter the password created during initial configuration of the Cisco MXE 3500.

> **Note**   If you upgraded your system from Cisco MXE 3500 Release 3.2 to Release 3.3, and have not synchronized the admin passwords, the default **ncpadmin** password is **idc**.

Step 3   Click on the Vocabulary tab.

Step 4   Click on **Upload Pulse Vocabulary**.

Figure 4-1   *Upload Pulse Vocabulary*

Step 5   Click on **Browse** to select the vocabulary file for upload.

Step 6   Click **Submit**.

Step 7   Click on the Administration tab. The Pulse Connect Appliance page is displayed.

Step 8   Click **Save**.

Figure 4-2   *Save Pulse Vocabulary*

Step 9   Click Configuration Update, and click on **Update Now**.

This updates Pulse video analytics with the new vocabulary list. If not, the Pulse video analytics is updated at the scheduled update.

Figure 4-3   *Update Pulse Vocabulary*

Step 10   Go to .

Configure the Speech Recognition Engine

A separate engagement with Cisco Advanced Services is required to configure the speech recognition engine used by Pulse video analytics. Please contact your Cisco representative for details.

Update the Document Repository and the Vocabulary List

- Update Document Repository, page 4-6
- Generate a New Vocabulary List, page 4-6
- Download and Merge Vocabulary Lists, page 4-6
- Upload the Merged Vocabulary List, page 4-7

Update Document Repository

Update the document repository when there are significant number of new documents created or added to your web domain.

Follow the steps detailed in Create a Document Repository, page 4-2.

Generate a New Vocabulary List

Follow the steps in Generate the Vocabulary List Using the AutoVocab Tool, page 4-4

The AutoVocab tool generates a new list which may include keywords already generated from the previous AutoVocab run.

After a new vocabulary list is created, follow the steps in Download and Merge Vocabulary Lists, page 4-6 to merge the old and new vocabulary lists.
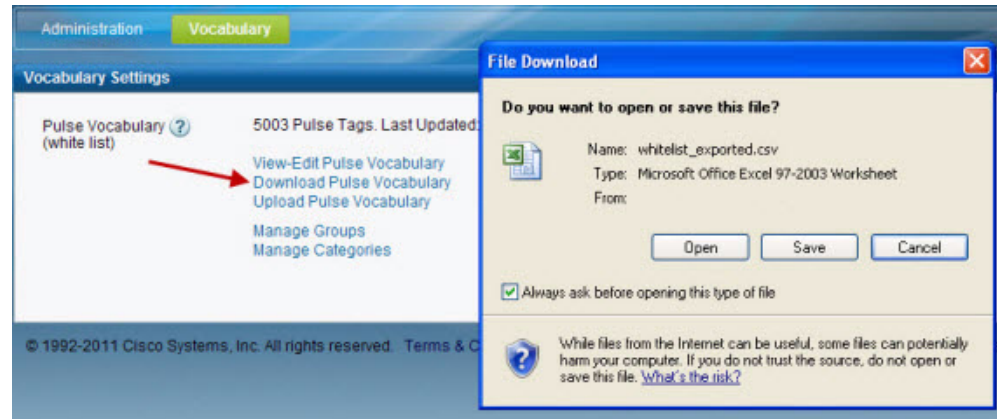
Download and Merge Vocabulary Lists

Step 1   Go to **http://**_mxe_IP_address_**/pulse/**, where _mxe_IP_address_ is the hostname or IP address for the Cisco MXE 3500.

Step 2   Login as **ncpadmin.** Enter the password created during initial configuration of the Cisco MXE 3500.

> ✎
>
> **Note**   If you upgraded your system from Cisco MXE 3500 Release 3.2 to Release 3.3, and have not synchronized the admin passwords, the default **ncpadmin** password is **idc**.

Step 3   Click on the Vocabulary tab.

Step 4   Click on **Download Pulse Vocabulary**.

Figure 4-4   Download Pulse Vocabulary

Step 5   Merge the new vocabulary list with the downloaded vocabulary list.

Tip   Use Microsoft Excel to merge, add, remove, and edit the vocabulary list.

Upload the Merged Vocabulary List

See Upload the Vocabulary List, page 4-4.

Speaker Identification

Pulse video analytics automatically identifies and labels speakers in a video and tags each speech segment. Speaker identification requires user input in Cisco Show and Share.

Each unique speaker is identified by a separate color bar in Cisco Show and Share. Each speech segment is identified by the speaker's unique color bar. The user selects a color bar to tag it with the speaker's name.

In subsequent videos, the corresponding speech segments are automatically labeled with the speaker name.

Cisco Show and Share Integration

For information on how to integrate the Cisco MXE 3500 with Cisco Show and Share, see *Integration Note for Configuring Cisco MXE 3500 Release 3.3 with Cisco Show and Share* on Cisco.com.

Checking Pulse Usage Hours

Step 1   SSH to *mxe_IP_address*, and login as **admin**.

Step 2   Select **System Command Prompt** and select **OK**.

Step 3   Enter **su -**.

Step 4   Enter **/opt/cisco/cli/queryindex**.

Step 5   (Perform this step, the first time only, to fix class not found exception.)
Enter **cp /opt/mgmt/lib/slf4j-jdk14-1.5.0.jar lib**.

Step 6   Enter **./usage_status.sh**.

```
Example Output
[root@mxe queryindex]# ./usage_status.sh
current total video duration = 30 hours 10 minutes 7 seconds
```