



Extracting and Preparing Metadata to Make Video Files Searchable

Meeting the Unique File Format and Delivery Requirements of Content Aggregators and Distributors

Table of Contents

- Executive Overview 2
- Creating Metadata for Videos 2
- Meeting the Evolving Requirements of Content Aggregators and Distributors 3
- Five Steps for Making Video Files Searchable 4
 - Step One: Determine which fields are used for searching 4
 - Step Two: Set up automated metadata extraction 4
 - Step Three: Set up manual annotation processes 4
 - Step Four: Convert and package the metadata 4
 - Step Five: Deliver the video and metadata to the searchable destination 4
- Using Telestream Solutions for Metadata Creation, Preparation, and Delivery 5
- Summary 6
- About Telestream 6

Executive Overview

Metadata is ‘data about data’ – or the information about a video file, audio clip, or Web page. Metadata can be used to facilitate the searching of video content either by individuals or corporations hosting their videos on their own websites, as well as for videos that are uploaded and hosted on video search engines and video sharing sites like Google, iTunes, YouTube, MySpace, and others.

Metadata is just a label on the video – it does not affect the video content itself. But once metadata is created or extracted from a video file, relevant keywords and descriptions can then be selected to drive effective search engine optimization (SEO). The metadata is one important aspect used by search engines to rank content in a search directory. Metadata is also used to generate short descriptions of the web page(s) or videos in the actual search results, which improves the volume and quality of traffic to a web site from the various search engines.

The creators and distributors of videos who submit files for mobile phones, podcasting, VOD, and web portals face several major challenges. They need to create relevant metadata and keywords out of the video so the video can be found in a searchable environment. After the metadata has been created or extracted, they need to deliver the video into the searchable repositories, in the right format, with the right metadata attached. This paper provides information on the methodology and solutions that enable video content creators to address these challenges.

Creating Metadata for Videos

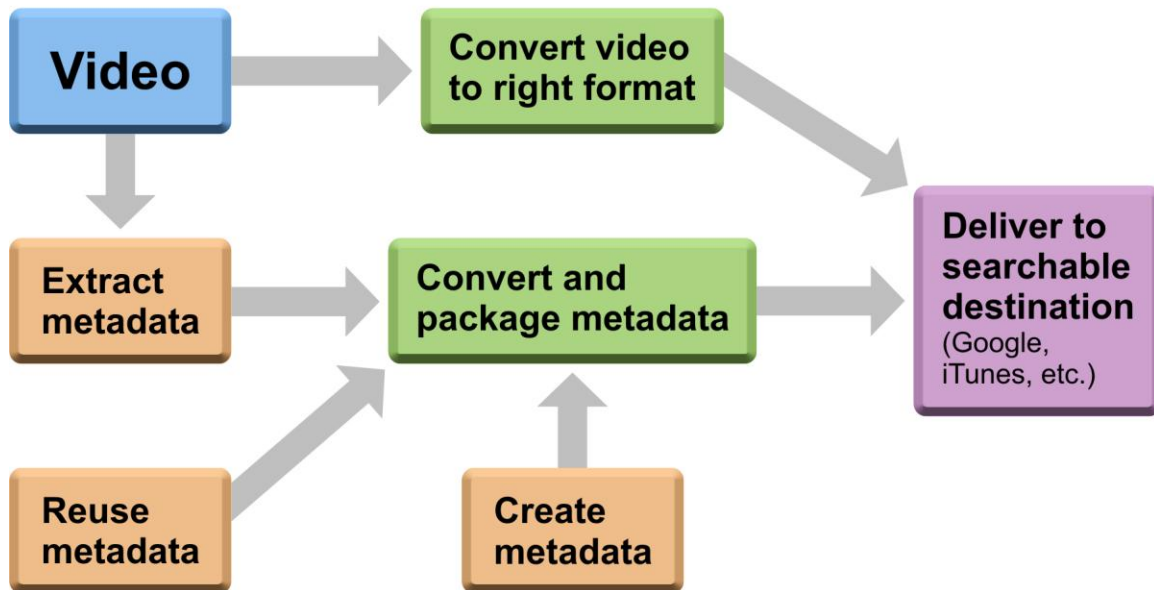
Creating metadata for text-based files is relatively easy, because text-based files are made up of words. There are many tools available today that will help content owners create metadata for these types of files. Videos present a much more formidable challenge. The old saying that “a picture is worth a thousand words” may be true, but unfortunately not for those tasked with creating metadata out of video files. Some videos may contain associated audio text and captions, others may only contain images. In either case, creating metadata from videos is a more complicated proposition than from files that contain primarily text.

There are several ways to extract and create metadata from a video:

- 1. Extract metadata from closed captions and other embedded video information from a source file.** An example would be when a person in a movie is talking about flying airplanes – those words can be converted to text. Operators can then search the extracted text manually, or through the use of automated software that will search for relevant keywords. These keywords can be extracted from this text file and later used for searching.
- 2. Re-use existing metadata, when it is present.** Some video formats, particularly those which are used for web streaming or editing, may already contain metadata such as descriptions of individual scenes. In some cases, it is possible to preserve and extract this information to facilitate searching.

3. **Create annotations or tagging** – another way to create metadata from a video is to have an operator manually type information as the video is being viewed. This can be done in a separate database, MS Word, or notepad file, or with automated software that can then link the created text to the video file.

Creating and Packaging Video Metadata



No matter which method is used, after creating or extracting metadata, that metadata (and the video itself) must then be converted and packaged into the appropriate formats specified by the recipients – typically a content aggregator or distributor.

Meeting the Evolving Requirements of Content Aggregators and Distributors

Each aggregator (e.g., iTunes, YouTube, etc.) has its own metadata file format and delivery requirements. To complicate matters even more, new video hosting sites emerge regularly, and existing aggregators may change their metadata requirements frequently as technology evolves and user requirements escalate.

Some aggregators require submitters to send the video, then the metadata files – others stipulate the reverse order. Some aggregators require that the video and metadata reside in the same file – compressed in a certain format with the right structure inside. Finally, many video formats support metadata attached directly to the file (in-file metadata), and some aggregators support this method of submission. In this final example, attaching the metadata to the video file has the advantage that the metadata cannot be lost by the aggregator or separated from the video when a user downloads the file to his or her computer instead of viewing it on the internet.

Many aggregators provide public specifications for their delivery schemes. For example, YouTube has several methods for content delivery that can be directly accessed from its website. Other aggregators, such as iTunes and many VOD aggregators, have special requirements that are only provided to content providers who are actually uploading content. If you plan to support uploading to an aggregator, it is worth reviewing their requirements as early as possible.

Five Steps for Making Video Files Searchable

There are five key steps content submitters can follow to facilitate the capture and creation of metadata, ultimately making video files more searchable:

Step One: Determine which fields are used for searching.

It is important to understand the requirements of the system you are delivering to. Each aggregator prioritizes and accepts specified fields. These fields can vary from one provider to the next. Therefore, it is important to identify the fields which will be used for Internet searching by the provider you plan to submit to. Metadata may include author, description, keywords, when to start playing, when to stop playing, episodes, etc. Of these, *description* and *keyword* fields are likely used by providers to make videos searchable on the Internet. However, each provider will have its own set of specific fields and rules that it follows. Check with your provider for more information.

Step Two: Set up automated metadata extraction.

Content creators can speed up metadata selection by creating a text file where operators can look at the captions and extract keywords. This means that operators don't have to watch every video and write out captions manually. Alternatively, keyword extraction software can be used on the captions to determine the best words for searching. Some software solutions can also extract video properties and read external XML files to automate the collection of metadata.

Step Three: Set up manual annotation processes.

Metadata can also be entered by hand. In a simplistic implementation, a person might watch each video and type keywords into a text file, MS Word document, or a spreadsheet. The challenge created by this process is that the video and the metadata are now in separate files. Fortunately, effective metadata data-entry systems exist which can provide one place to watch the video and enter notes which are attached to the video within the system. These solutions can eliminate the complexity of having separate documents or spreadsheets with metadata and trying to match them up with the video file later. They also ensure that the metadata always stays with the video.

Step Four: Convert and package the metadata.

After the metadata has been extracted and/or annotated, it must be formatted according to the aggregator's requirements.

Step Five: Deliver the video and metadata to the searchable destination.

In the final step, the video content is transcoded to the appropriate format for each destination, and then both the metadata and the video are delivered. This is where flexible

software solutions enable content owners, creators, aggregators, and distributors to choose the functionality and feature sets to meet any digital media workflow requirement. Workflows typically include what video file format to create, whether to FTP video files or use a different delivery protocol, and the ability to select processes which instruct when to deliver either the metadata first, or the video first.

Using Telestream Solutions for Metadata Creation, Preparation, and Delivery

Video content creators and submitters are now seeking packaged solutions to assist them with the process of combining metadata with media to meet the unique file format and delivery requirements of a growing number of content aggregators and distributors. Telestream's products include tools and capabilities that facilitate and automate the processes for extracting, annotating, formatting, and delivering video and metadata.

Telestream's FlipFactory[®] workflow automation software products simplify the process of extracting information for the creation of metadata. The software is installed on a server, and it watches for video files that are placed in one or more hot folders. When files appear, it can automatically begin to process the files (or if a user prefers, a file can be manually submitted). Each file is sent to a *factory* whose assigned set of rules can include what format to create, analyses to be performed, metadata packaging guidelines, and where to deliver the file. These rules are chosen or built by the FlipFactory administrator.

In the simplest case, the video content owner sets up the rules, video files are dropped into a hot folder, and the *factory* does all of the conversion, packaging, and sending automatically. The administrator can build in steps requiring the system to wait for user input, allowing operators to review videos and annotate metadata.

Included with FlipFactory is Telestream's exclusive MetaFlip[®] technology, which allows users to automatically flip (transform and deliver) metadata along with video and audio files to meet the data requirements of content distribution service providers. MetaFlip enables easy, accurate data entry and review through automatically generated data entry forms. This simplifies the workflow and quality assurance process to deliver and package metadata and video in the format specified by the delivery destination.

MetaFlip currently supports iTunes RSS requirements and CableLabs VOD workflows, with support for other major content websites being added. MetaFlip is included as a standard feature in all FlipFactory products. With MetaFlip, metadata files can be:

- **Generated automatically by transforming metadata sources.** MetaFlip picks up XML files containing metadata and translates them to simultaneously create any number of output metadata files in parallel with their corresponding video files.
- **Generated automatically from data sources.** Using the FlipFactory SDK, metadata can be obtained from virtually any source to create the metadata files.

- **Created supporting existing standards or unique formats.** Using XSL stylesheets, MetaFlip offers a truly open metadata platform allowing *virtually any output metadata file format to be supported*. Simply provide the appropriate stylesheet, and MetaFlip will start generating the files. Standard metadata file support includes RSS feeds, iTunes, and VOD XML formats.
- **Created manually using data entry forms.** MetaFlip presents a form to operators allowing them to perform data entry for video files within the FlipFactory system. Default values, drop-down options, and form structure can be configured by the FlipFactory MetaFlip administrator.
- **Generated in one pass.** A single data source can be transformed into multiple output files, each with unique format and delivery specifications.
- **Delivered with the video file.** Using FlipFactory's built-in delivery services, metadata files are delivered with the video files through common protocols, directly to content aggregators.

Using FlipFactory with MetaFlip automates metadata workflows with a single, scalable solution.

Summary

Video objects are very hard to search on the web. Although many search engines are becoming more multimedia aware and allow users to search video keywords and descriptions, the more information the content creator can attach to the video file, the greater the likelihood that viewers will be able to locate and view the video. This results in higher visibility and possibly increased income for the video creators.

Telestream provides a robust set of video software tools that both allow for metadata information to be embedded within the files themselves and/or prompt a user to enter metadata before the file is submitted to the aggregator. With Telestream FlipFactory and MetaFlip technology, the task of creating metadata for videos has finally gotten much easier.

About Telestream

Telestream products make it possible to get video content to any audience regardless of how it is created, distributed, or viewed. Throughout the entire digital media lifecycle, from capture to viewing, for consumers through high-end professionals, Telestream products range from components to fully automated workflow systems. Telestream enables users in a broad range of business environments to leverage the value of their video content. Telestream corporate headquarters are located in Nevada City, California and its team of video experts are located in France, Germany, Sweden, UK and US. The company is privately held.

For more information on Telestream products and services, visit www.telestream.net or contact us at +1 530 470 1300.