

EXIFToolParser

Tika supports EXIFTool now through the External parser. Read on to find out how to use it.

Download and install EXIFTool

[EXIFTool](#) is a wonderful tool that reads videos, images, audio and other media files and that extracts EXIF metadata from them. If you're lucky, you can install EXIFTool with the following commands.

On Mac

```
brew install exiftool
```

On Linux (CentOS)

```
sudo yum install perl-Image-ExifTool
```

To verify that EXIFTool works correctly, run:

```
exiftool -ver
```

which should output something like: 9.72

Using EXIFTool with Tika

To use EXIFTool you'll need a custom Tika config that will override Tika's default MP4 parser (if you are dealing with MP4 files). You can do so by creating a file such as the one below:

```
<properties>
  <parsers>
    <parser class="org.apache.tika.parser.DefaultParser">
    </parser>
    <parser class="org.apache.tika.parser.mp4.MP4Parser">
      <mime-exclude>video/mp4</mime-exclude>
    </parser>
    <parser class="org.apache.tika.parser.external.CompositeExternalParser">
      <mime>video/mp4</mime>
    </parser>
  </parsers>
</properties>
```

Note that this config file initializes the [DefaultParser](#) a [CompositeParser](#), and the [CompositeExternalParser](#), and the MP4Parser. For the MP4Parser, it uses a new directive, mime-exclude, to exclude that parser from the video/mp4 type, and then to declare that [CompositeExternalParser](#) will support video/mp4. Since EXIFTool is an [ExternalParser](#) this configuration will make sure it gets called.

Once you have the config file made above, save it as a file, e.g., `exif-tika-config.xml` in the current directory. Then to call Tika, you can use Tika-App and/or Tika Server.

Using Tika-App

Use the following command on a file, e.g., `spaghetti-to-sushi.mp4`:

```
java -Dtika.config=exif-tika-config.xml -classpath tika-app/target/tika-app-1.9-SNAPSHOT.jar org.apache.tika.cli.TikaCLI -m spaghetti-to-sushi.mp4
```

This should output:

```
Audio Bits Per Sample: 16
Audio Channels: 2
Audio Format: mp4a
Audio Sample Rate: 22050
```

Average Bitrate: 0
Avg Bitrate: 1.26 Mbps
Balance: 0
Bit Depth: 24
Buffer Size: 0
Compatible Brands: mp41
Compressor ID: avc1
Compressor Name: h264
Content Create Date: created.with.SUPER(C).v2006.19
Content Create Date (ja): created.with.SUPER(C).v2006.19
Content-Length: 353985630
Content-Type: video/mp4
Create Date: 2006:12:17 18:50:47
Current Time: 0 s
Duration: 0:37:19
Elementary Stream Track: 201 101
ExifTool Version Number: 9.72
File Access Date/Time: 2015:05:25 21:18:08-07:00
File Inode Change Date/Time: 2014:09:26 20:32:27-07:00
File Modification Date/Time: 2011:07:28 13:01:54-07:00
File Name: spaghetti-to-sushi.mp4
File Permissions: rwxr-xr-x
File Size: 338 MB
File Type: MP4
Graphics Mode: srcCopy
Handler Description: GPAC MPEG-4 BIFS Handler
Handler Type: Metadata
Handler Vendor ID: Apple
Image Height: 480
Image Size: 640x480
Image Width: 640
MIME Type: video
Major Brand: MP4 v2 [ISO 14496-14]
Matrix Structure: 1 0 0 0 1 0 0 0 1
Max Bitrate: 0
Media Create Date: 2006:12:16 20:07:48
Media Duration: 1.00 s
Media Header Version: 0
Media Language Code: und
Media Modify Date: 2006:12:16 20:07:48
Media Time Scale: 90000
Minor Version: 0.0.1
Modify Date: 2006:12:17 18:50:47
Movie Data Offset: 473003
Movie Data Size: 353512586
Movie Header Version: 0
Next Track ID: 201
Op Color: 0 0 0
Other Format: mp4s
Poster Time: 0 s
Preferred Rate: 1
Preferred Volume: 100.00
Preview Duration: 0 s
Preview Time: 0 s
Rotation: 0
Selection Duration: 0 s
Selection Time: 0 s
Source Image Height: 480
Source Image Width: 720
Time Scale: 90000
Title: From Spaghetti to Sushi.mpeg
Title (ja): From Spaghetti to Sushi.mpeg
Track Create Date: 2006:12:17 18:50:47
Track Duration: 0:37:19
Track Header Version: 0
Track ID: 201
Track Layer: 0
Track Modify Date: 2006:12:16 20:07:48
Track Volume: 0.00
Vendor ID: FFmpeg
Video Frame Rate: 25

```
X Resolution: 72
X-Parsed-By: org.apache.tika.parser.CompositeParser
X-Parsed-By: org.apache.tika.parser.external.CompositeExternalParser
X-Parsed-By: org.apache.tika.parser.external.ExternalParser
Y Resolution: 72
resourceName: spaghetti-to-sushi.mp4
```

Using Tika Server

You can also use Tika-Server. First, start it up:

```
java -Dtika.config=exif-tika-config.xml -classpath tika-server/target/tika-server-1.9-SNAPSHOT.jar org.apache.tika.server.TikaServerCli
```

Now, PUT a file to it, e.g., spaghetti-to-sushi.mp4:

```
curl -T $HOME/Movies/spaghetti-to-sushi.mp4 -H "Content-Disposition: attachment;filename=spaghetti-to-sushi.mp4" http://localhost:9998/rmeta
```

Which should return:

```
[
  {
    "Audio Bits Per Sample": "16",
    "Audio Channels": "2",
    "Audio Format": "mp4a",
    "Audio Sample Rate": "22050",
    "Average Bitrate": "0",
    "Avg Bitrate": "1.26 Mbps",
    "Balance": "0",
    "Bit Depth": "24",
    "Buffer Size": "0",
    "Compatible Brands": "mp41",
    "Compressor ID": "avc1",
    "Compressor Name": "h264",
    "Content Create Date": "created.with.SUPER(C).v2006.19",
    "Content Create Date (ja)": "created.with.SUPER(C).v2006.19",
    "Content-Type": "video/mp4",
    "Create Date": "2006:12:17 18:50:47",
    "Current Time": "0 s",
    "Duration": "0:37:19",
    "Elementary Stream Track": "201 101",
    "ExifTool Version Number": "9.72",
    "File Access Date/Time": "2015:05:25 21:20:47-07:00",
    "File Inode Change Date/Time": "2015:05:25 21:20:46-07:00",
    "File Modification Date/Time": "2015:05:25 21:20:46-07:00",
    "File Name": "apache-tika-3052147227532168299.tmp",
    "File Permissions": "rw-r--r--",
    "File Size": "338 MB",
    "File Type": "MP4",
    "Graphics Mode": "srcCopy",
    "Handler Description": "GPAC MPEG-4 BIFS Handler",
    "Handler Type": "Metadata",
    "Handler Vendor ID": "Apple",
    "Image Height": "480",
    "Image Size": "640x480",
    "Image Width": "640",
    "MIME Type": "video",
    "Major Brand": "MP4 v2 [ISO 14496-14]",
    "Matrix Structure": "1 0 0 0 1 0 0 0 1",
    "Max Bitrate": "0",
    "Media Create Date": "2006:12:16 20:07:48",
    "Media Duration": "1.00 s",
    "Media Header Version": "0",
    "Media Language Code": "und",
    "Media Modify Date": "2006:12:16 20:07:48",
```

```
"Media Time Scale":"900000",
"Minor Version":"0.0.1",
"Modify Date":"2006:12:17 18:50:47",
"Movie Data Offset":"473003",
"Movie Data Size":"353512586",
"Movie Header Version":"0",
"Next Track ID":"201",
"Op Color":"0 0 0",
"Other Format":"mp4s",
"Poster Time":"0 s",
"Preferred Rate":"1",
"Preferred Volume":"100.00",
"Preview Duration":"0 s",
"Preview Time":"0 s",
"Rotation":"0",
"Selection Duration":"0 s",
"Selection Time":"0 s",
"Source Image Height":"480",
"Source Image Width":"720",
"Time Scale":"900000",
"Title":"From Spaghetti to Sushi.mpeg",
"Title (ja)":"From Spaghetti to Sushi.mpeg",
"Track Create Date":"2006:12:17 18:50:47",
"Track Duration":"0:37:19",
"Track Header Version":"0",
"Track ID":"201",
"Track Layer":"0",
"Track Modify Date":"2006:12:16 20:07:48",
"Track Volume":"0.00",
"Vendor ID":"FFmpeg",
"Video Frame Rate":"25",
"X Resolution":"72",
"X-Parsed-By":[
    "org.apache.tika.parser.CompositeParser",
    "org.apache.tika.parser.external.CompositeExternalParser",
    "org.apache.tika.parser.external.ExternalParser"
],
"X-TIKA:parse_time_millis":"3638",
"Y Resolution":"72",
"resourceName":"spaghetti-to-sushi.mp4"
```

```
}
```

```
]
```