

FFMPEGParser

Tika now support FFMPEG extraction. Read on below to see how.

Install FFMPEG

If you're lucky, the following should install FFMPEG.

```
brew install ffmpeg
```

You can test to see if FFMPEG is installed by typing:

```
ffmpeg -version
```

You should see something resembling:

```
ffmpeg version 2.3.3 Copyright (c) 2000-2014 the FFmpeg developers
built on Jan  8 2015 14:52:39 with Apple LLVM version 6.0 (clang-600.0.56) (based on LLVM 3.5svn)
configuration: --prefix=/usr/local/Cellar/ffmpeg/2.3.3 --enable-shared --enable-pthreads --enable-gpl --enable-
version3 --enable-nonfree --enable-hardcoded-tables --enable-avresample --enable-vda --cc=clang --host-cflags=
--host-ldflags= --enable-libx264 --enable-libfaac --enable-libmp3lame --enable-libxvid
libavutil      52. 92.100 / 52. 92.100
libavcodec     55. 69.100 / 55. 69.100
libavformat    55. 48.100 / 55. 48.100
libavdevice    55. 13.102 / 55. 13.102
libavfilter    4. 11.100 / 4. 11.100
libavresample  1.  3.  0 / 1.  3.  0
libswscale     2.  6.100 / 2.  6.100
libswresample  0. 19.100 / 0. 19.100
libpostproc   52.  3.100 / 52.  3.100
```

To use FFMPEG in Tika, then you simply use Tika-App and/or the Tika-REST server on a video file. Read on below to see how.

Using Tika App

You can use Tika app once FFMPEG is installed to parse a video file. For example:

```
java -classpath tika-app/target/tika-app-1.9-SNAPSHOT.jar org.apache.tika.cli.TikaCLI -m SPOT11_000001\ 15.AVI
```

Which should produce the following output:

```
Content-Length: 312559634
Content-Type: video/x-msvideo
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.external.CompositeExternalParser
X-Parsed-By: org.apache.tika.parser.external.ExternalParser
encoder: ankarec
resourceName: SPOT11_000001 15.AVI
videoResolution: 720x480
xmpDM:audioChannelType: 1
xmpDM:audioCompressor: pcm_s16le ([1][0][0][0] / 0x0001)
xmpDM:audioSampleRate: 8000
xmpDM:duration: 00:05:35.92
xmpDM:fileDataRate: 7443 kb/s
xmpDM:videoColorSpace: yuvj420p(pc, bt470bg)
xmpDM:videoCompressor: mjpeg (MJPEG / 0x47504A4D)
xmpDM:videoFrameRate: 25
```

Here is a test on another file:

```
java -classpath tika-app/target/tika-app-1.9-SNAPSHOT.jar org.apache.tika.cli.TikaCLI -m WOW_MR_T.avi
```

Which should produce the following output:

```
Content-Length: 8932074
Content-Type: video/x-msvideo
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.external.CompositeExternalParser
X-Parsed-By: org.apache.tika.parser.external.ExternalParser
resourceName: WOW_MR_T.avi
xmpDM:audioCompressor: mp3 (U[0][0][0] / 0x0055)
xmpDM:audioSampleRate: 48000
xmpDM:duration: 00:00:33.03
xmpDM:fileDataRate: 2163 kb/s
xmpDM:videoCompressor: mpeg4 (DX50 / 0x30355844)
```

Using Tika Server

Start Tika server using the following command:

```
java -jar tika-server/target/tika-server-1.9-SNAPSHOT.jar
```

Then, issue a cURL command to post a video to Tika Server for FFMPEG to parse:

```
curl -T WOW_MR_T.avi -H "Content-Disposition: attachment; filename=WOW_MR_T.avi" http://localhost:9998/rmeta
```

Which should return:

```
[
  {
    "Content-Type": "video/x-msvideo",
    "X-Parsed-By": [
      "org.apache.tika.parser.DefaultParser",
      "org.apache.tika.parser.external.CompositeExternalParser",
      "org.apache.tika.parser.external.ExternalParser"
    ],
    "X-TIKA:parse_time_millis": "219",
    "resourceName": "WOW_MR_T.avi",
    "xmpDM:audioCompressor": "mp3 (U[0][0][0] / 0x0055)",
    "xmpDM:audioSampleRate": "48000",
    "xmpDM:duration": "00:00:33.03",
    "xmpDM:fileDataRate": "2163 kb/s",
    "xmpDM:videoCompressor": "mpeg4 (DX50 / 0x30355844)"
  }
]
```